

# LARGE DEVIATION FOR DIFFUSIONS AND HAMILTON–JACOBI EQUATION IN HILBERT SPACES

BY JIN FENG

*University of Massachusetts–Amherst*

Large deviation for Markov processes can be studied by Hamilton–Jacobi equation techniques. The method of proof involves three steps: First, we apply a nonlinear transform to generators of the Markov processes, and verify that limit of the transformed generators exists. Such limit induces a Hamilton–Jacobi equation. Second, we show that a strong form of uniqueness (the comparison principle) holds for the limit equation. Finally, we verify an exponential compact containment estimate. The large deviation principle then follows from the above three verifications.

This paper illustrates such a method applied to a class of Hilbert-space-valued small diffusion processes. The examples include stochastically perturbed Allen–Cahn, Cahn–Hilliard PDEs and a one-dimensional quasilinear PDE with a viscosity term. We prove the comparison principle using a variant of the Tataru method. We also discuss different notions of viscosity solution in infinite dimensions in such context.

**1. Introduction.** We are interested in large deviation for small randomly perturbed diffusion processes in a Hilbert state space  $E$ . When  $E = R^d$ , this is known as the Freidlin and Wentzell theory [23]. The proofs in [23] rely upon the Girsanov transformations. The idea is to estimate probability of an atypical, large deviant event under the given probability law through a change of measure, so that the event becomes most probable under the new law. Such technique is also repeatedly used in the Donsker and Varadhan theory [13] regarding occupation measures, which is another kind of large deviation concerning ergodic phenomena instead of small random perturbations.

There exists a different approach to the above mentioned large deviation problems. In the late 1970s, Fleming [19] introduced a logarithmic trans-

---

Received January 2002; revised January 2005.

*AMS 2000 subject classifications.* Primary 60F10; secondary 60J25, 49L25, 60G99.

*Key words and phrases.* Large deviation, stochastic evolution equation in Hilbert space, viscosity solution of Hamilton–Jacobi equations.

This is an electronic reprint of the original article published by the Institute of Mathematical Statistics in *The Annals of Probability*, 2006, Vol. 34, No. 1, 321–385. This reprint differs from the original in pagination and typographic detail.

form to generators of Markov processes, giving exit probabilities an optimal control interpretation. This observation allowed us to characterize the large deviation convergence for exit probabilities as convergence of solutions for a sequence of Hamilton–Jacobi equations. Later, Evans and Ishii [16], Fleming and Souganidis [21, 22], among others, applied the theory of viscosity solution to this context, enabling the approach to cover a wider variety of examples. In particular, this includes  $R^d$ -valued diffusions with vanishing stochastic terms. During early developments of this approach, the applicable settings and structural conditions required were relatively restrictive as compared to the Girsanov transformation approach. However, this can be fixed by refining techniques on the viscosity solution techniques and on the large deviation theory. Feng and Kurtz [18] recently carried out such a program which expands the theory.

The general setting in [18] allows the state space  $E$  to be a metric space. One of the key technical conditions assumed is a strong form of uniqueness (i.e., the comparison principle, Definition 1.15) for a limit Hamilton–Jacobi type equation. For small perturbation type large deviations, such equation is usually a first-order nonlinear partial differential equation. When  $E = R^d$ , or a subset of it, the comparison principle can usually be verified by well-known criteria in PDE theory. Using these techniques, [18] treats both the classical Freidlin–Wentzell theory and the Donsker–Varadhan theory within one framework using the generator convergence approach.

When we study large deviation for stochastic PDEs or interacting particles, we usually encounter function- or measure-valued state space. Comparison principles of these types, however, are much less well understood. On the one hand, there exists an extensive PDE literature regarding first-order Hamilton–Jacobi equations in Hilbert/Banach spaces (e.g., [7, 8, 31, 32]). On the other hand, the operators derived in the large deviation context frequently exhibit subtle differences relative to those studied in the PDE literature. Indeed, in the case of applications to interacting particle systems, it is more natural to consider the state space as the space of probability measures, rather than a Banach space.

We restrict attention to Hilbert-space-valued diffusions only in this paper.

**1.1. Background.** We consider the large deviation for Hilbert-space-valued diffusions with a possibly nonlinear drift term. To outline the approach and identify difficulties ahead of us, first, we review a general result (adapted to the situation of this paper) developed in [18].

Let  $X_n$ ,  $n = 1, 2, \dots$ , be a sequence of metric-space  $S$ -valued random variables. Varadhan and Brycs (e.g. Theorems 4.3.1 and 4.4.2 in [12]) discovered the following moment characterization of large deviation convergence.

**PROPOSITION 1.1.**

- (a) Suppose  $\{X_n\}$  satisfies the large deviation principle (Definition 1.17) with a good rate function  $I$ . Then for each  $f \in C_b(S)$  (bounded continuous functions on  $S$ ), if we define  $\Lambda_n(f) = n^{-1} \log E[\exp\{nf(X_n)\}]$ ,

$$(1.1) \quad \lim_{n \rightarrow +\infty} \Lambda_n(f) = \lim_{n \rightarrow +\infty} \frac{1}{n} \log E[e^{nf(X_n)}] = \sup_{x \in S} \{f(x) - I(x)\} = \Lambda(f).$$

- (b) Suppose that  $\{X_n\}$  is exponentially tight (Definition 1.17) and that the limit (1.1) exists for each  $f \in C_b(S)$ . Then  $\{X_n\}$  satisfies the large deviation with good rate function

$$(1.2) \quad I(x) = \sup_{f \in C_b(S)} (f(x) - \Lambda(f)).$$

See Theorems 4.3.1 and 4.4.2 in [12].

The main result in [18] can be viewed as a process version of the above theorem, expressed at an infinitesimal level.

To explain the result, we proceed informally first. Let  $\{X_n(t), 0 \leq t < +\infty; n = 1, 2, \dots\}$  denote a sequence of metric-space  $E$ -valued Markov processes. For simplicity, we assume the trajectories are continuous. By the Markov property and continuity of the trajectories, we expect the large deviation of  $\{X_n\}$  follows from that of the transition probability measures  $\{P(X_n(t) \in dy | X_n(0) = x)\}$  for all  $x \in E$  and  $t \geq 0$ . By Proposition 1.1, we also expect this to be implied by convergence of the functionals

$$V_n(t)f(x) \equiv \frac{1}{n} \log E[e^{nf(X_n(t))} | X_n(0) = x] \rightarrow V(t)f(x) \quad \text{for some } V(t)f,$$

where  $f \in D \subset C_b(E)$ , and  $D$  is sufficiently dense in  $C_b(E)$  in appropriate sense.

It turns out that, by the Markov property,  $V_n$  forms a nonlinear operator semigroup

$$V_n(s)V_n(t) = V_n(t+s), \quad s, t \geq 0.$$

Hence  $\{V(t) : t \geq 0\}$ , viewed as a collection of operators acting on functions, should form a semigroup as well. We identify the generator of  $V_n$  next:

$$\begin{aligned} H_n f(x) &= \lim_{t \rightarrow 0+} \frac{1}{t} (V_n(t)f(x) - V_n(0)f(x)) \\ &= \lim_{t \rightarrow 0+} \frac{1}{t} \frac{1}{n} \log E[e^{n(f(X_n(t)) - f(x))} | X_n(0) = x] \\ &= \frac{1}{n} e^{-nf(x)} A_n e^{nf(x)}, \end{aligned}$$

where  $A_n$  is the generator for the process  $X_n$

$$A_n g(x) = \lim_{t \rightarrow 0+} \frac{1}{t} \log E[g(X_n(t)) - g(x) | X_n(0) = x].$$

The above transformation from  $A_n$  to  $H_n$  is essentially the logarithmic transform by Fleming [19]. If we denote by  $H$  the generator for  $V$ , then we expect generator convergence  $H_n \rightarrow H$  will imply semigroup convergence  $V_n \rightarrow V$ , which is suggested by the semigroup generation theorem:

$$V_n(t)h = \lim_{k \rightarrow +\infty} \left( I - \frac{t}{k} H_n \right)^{-k} h$$

and

$$(1.3) \quad V(t)h = \lim_{k \rightarrow +\infty} \left( I - \frac{t}{k} H \right)^{-k} h.$$

Going backward in the reasoning, modulo regularity conditions, we expect convergence  $H_n \rightarrow H$  will give the large deviation of  $\{X_n\}$ .

There are practical problems if we want to rigorously apply the above program to examples. First, the formula in (1.3) requires

$$(1.4) \quad (I - \alpha H)f = h$$

to hold in the classical sense for all  $h \in D$  and  $\alpha > 0$  (if  $H$  is dissipative, then such  $f$  is also unique). However, this is extremely hard to verify for most examples. Therefore we are forced to modify the above formulation by using a type of weak solution called the *viscosity solution* (Definition 1.14). By weakening the type of solution needed for (1.4), we have to require a strong form of uniqueness condition known as the *comparison principle* (Definition 1.15). Informally, this principle states that, if upper semicontinuous  $\bar{f}$  and lower semicontinuous  $\underline{f}$  satisfy

$$(I - \alpha H)\bar{f} \leq h \quad \text{and} \quad (I - \alpha H)\underline{f} \geq h,$$

then  $\bar{f} \leq \underline{f}$ . The  $\bar{f}$  and  $\underline{f}$  are called, respectively, a subsolution and a supersolution. Existence of sub- and supersolutions, in the large deviation context here, can be constructed by generalizing a procedure due to Barles and Perthame [2, 3]. When  $E$  is noncompact, in order for such argument to go through, we need another crucial condition (Condition 2.2) for the transition probabilities. Such condition requires the processes to be concentrated on a compact subset of the state space with high probability.

Note that the above formulation is only based on inequalities for sub- and supersolutions. This provides an opportunity for further relaxations on conditions. We can introduce two more operators:  $H_0, H_1$  so that  $Hf \leq H_0f$  and  $Hf \geq H_1f$  for all  $f \in \mathcal{D}(H) \cap \mathcal{D}(H_i)$ . Then

$$(I - \alpha H_0)\bar{f} \leq h \quad \text{and} \quad (I - \alpha H_1)\underline{f} \geq h.$$

Suppose that the comparison principle still holds for the above two “inequalities” (i.e.,  $\bar{f} \leq \underline{f}$ ). The construction of  $\underline{f}, \bar{f}$  by the Barles–Perthame

procedure then reveals that  $\overline{f} = \underline{f} = f \in C_b(E)$ . Hence, each  $h$  uniquely corresponds to an  $f \in C_b(E)$ , and we can denote such correspondence by  $f = R_\alpha h$ . Consequently, at least formally,  $R_\alpha = (I - \alpha H)^{-1}$ . In other words,  $H_0, H_1$  implicitly determine  $H$  through its resolvent, and  $V(t)h = \lim_n R_{t/n}^n h \in C_b(E)$ . We can now completely avoid using  $H$  in the above program by replacing condition  $H_n \rightarrow H$  by: for each  $f \in \mathcal{D}(H_i)$ ,

$$H_1 f \leq \liminf_n H_n f_n, \quad \limsup_n H_n f_n \leq H_0 f \quad \text{some } f_n \rightarrow f.$$

The above generalization is useful for applications where  $E$  is infinite dimensional. We illustrate this next.

In general, the comparison principle proof relies upon test functions which behave like distance functions. For instance, in the case  $E = R^d$ , these test functions take the form  $f(x) = (\mu/2)|x - y|^2$ ,  $\mu \in R$  (see [5]). For  $H_n$ 's which are differential operators, there is no difficulty to include such functions in the domain. Furthermore, identifying  $Hf$  as a limit of  $H_n f$  is usually straightforward.

However, the situation becomes tricky when  $E$  is infinite dimensional (e.g., Examples 1.2, 1.5 and 1.8). For instance, let  $E = L^2(\mathcal{O})$  and  $\mathcal{O} = [0, 1]$  with periodic boundary, and

$$H_n f(x) = \langle \Delta x, Df(x) \rangle + \frac{1}{2} \|Df(x)\|^2 + o_f(1),$$

where  $o_f$  may depend on  $f$ . See (1.33) for definition of  $Df$ . We expect

$$Hf(x) = \langle \Delta x, Df(x) \rangle + \frac{1}{2} \|Df(x)\|^2.$$

But then, even for

$$(1.5) \quad f(x) = (\mu/2)\|x - y\|^2, \quad \mu \in R,$$

where  $y$  is arbitrarily smooth,

$$\lim_{x_n \rightarrow x} H_n f(x_n) \neq Hf(x).$$

Note that in this case,  $Df(x) = \mu(x - y)$  and  $\langle \Delta x, Df(x) \rangle$  is well defined as a function taking value in extended reals

$$\langle \Delta x, Df(x) \rangle = \mu \langle \Delta x, x - y \rangle = -\mu \|\nabla x\|^2 + \mu \langle \nabla x, \nabla y \rangle.$$

Assuming  $\mu > 0$ , by lower semicontinuity of  $\|\nabla x\|$ , we can however obtain  $\limsup_{x_n \rightarrow x} H_n f(x_n) \leq Hf(x)$ . Similarly, by reversing the inequality, we can verify a lower bound estimate for the case  $\mu < 0$ .

More generally, if the above  $\Delta$  is replaced by a general nonlinear dissipative operator  $C$  [assuming the domain of  $C$  is not the entire  $E$ ,  $\mathcal{D}(C) \neq E$ ], integration by parts may not even make sense any more. Consequently

$$Hf(x) = \langle Cx, Df(x) \rangle + \frac{1}{2} \|Df(x)\|^2$$

does not make sense for all  $x \in E$ , even if  $y \in \mathcal{D}(C)$ . Note that, if  $C$  is dissipative,

$$\mu \langle Cx, x - y \rangle \leq \mu \langle Cy, x - y \rangle \quad \forall x, y \in \mathcal{D}(C), \mu > 0.$$

The right-hand side of the above is continuous in  $x$ , and can be extended to all  $x \in E$  easily. Hence at least for  $f$  of the form (1.5) with  $\mu > 0$ , if  $H_n f(x_n)$  is defined, then it can be estimated from above by

$$\begin{aligned} \limsup_{x_n \rightarrow x} H_n f(x_n) &\leq \lim_n \mu \langle Cy, x_n - y \rangle + \frac{1}{2} \|Df(x_n)\|^2 + o_f(1) \\ &= \mu \langle Cy, x - y \rangle + \frac{1}{2} \|Df(x)\|^2 \equiv H_0 f(x). \end{aligned}$$

Furthermore, such  $H_0 f \in C(E)$  and is everywhere well defined. By the arbitrariness of  $y \in \mathcal{D}(C)$ , we hope such  $H_0$  provides a sharp estimate on the asymptotics of  $H_n$ 's. Similarly, we can also estimate  $H_n f$  from below by some  $H_1 f$ , if  $\mu < 0$ .

We call (1.4) a Hamilton–Jacobi equation, because of its connection with the optimal control problem. By the Markov property on the processes  $X_n$ , the  $A_n$ 's satisfy the maximum principle. The  $H_n$ 's, obtained as a transform of the  $A_n$ 's, also satisfy a nonlinear maximum principle. So does the limiting  $H$  (and frequently,  $H_0, H_1$ ). Using this property, the following variational representation of  $H_n$  can usually be proved [17]:

$$H_n f(x) = \sup_{u \in U} (B_n f(x, u) - L_n(x, u)),$$

where  $U$  is some auxiliary metric space, and for each  $u$  fixed,  $B_n f(\cdot, u)$  is a linear operator satisfying the maximum principle in  $C_b(E)$ ,  $L_n$  is a lower semicontinuous bivariate function. In the limit,  $H$  is supposed to have a similar structure. This is known as the Nisio representation of generator for Hamiltonian operator  $H$  in optimal control theory [25]. Based upon such representation, [18] proved theorems ensuring a simpler, variational representation of the rate function in an “action integral” form.

1.2. *Basic setup.* Let Hilbert-space-valued diffusion processes

$$(1.6) \quad dX_n(t) = \hat{C}_n X_n(t) dt + \frac{1}{\sqrt{n}} B_n(X_n(t)) dW(t),$$

where  $W$  is a cylindrical Wiener process [see (1.28)] on a separable real Hilbert space  $U_0$ ,  $E$  is another separable real Hilbert space,  $\hat{C}_n - \omega I$  is an  $m$ -dissipative (possibly) nonlinear operator on  $E$  for some  $\omega > 0$ : that is,

$$\langle \hat{C}_n x - \hat{C}_n y, x - y \rangle \leq \omega \|x - y\|^2 \quad \forall x, y \in \mathcal{D}(\hat{C}_n),$$

and the range of  $I - \alpha \hat{C}_n$  satisfies

$$\mathcal{R}(I - \alpha \hat{C}_n) = E \quad \forall \alpha > 0.$$

We also assume that  $0 \in \mathcal{D}(\hat{C}_n)$ , and  $B_n(x): U_0 \rightarrow E$  is a Hilbert–Schmidt operator for each  $x \in E$  fixed. More conditions are needed in order to make sense of the solution and large deviation result of (1.6); we delay them until the statement of the respective theorem. To simplify the presentation, we will actually deal with another form of the above equation:

$$(1.7) \quad dX_n(t) = C_n X_n(t) dt + F_n(X_n(t)) dt + \frac{1}{\sqrt{n}} B_n(X_n(t)) dW(t),$$

where  $C_n$  is  $m$ -dissipative,  $C_n 0 = 0$  and  $F_n(x): E \rightarrow E$  is globally Lipschitz in  $x$ . To rewrite (1.6) into the form of (1.7), we take

$$C_n x = \hat{C}_n x - \hat{C}_n 0 - \omega x, \quad F_n(x) = \hat{C}_n 0 + \omega x.$$

We provide three examples.

EXAMPLE 1.2 (Stochastic Allen–Cahn equation). Let  $\mathcal{O} = [0, 1]^d$ ,  $d = 1, 2, 3, \dots$ , with periodic boundary condition; we associate  $L^2(\mathcal{O})$  with the usual inner product

$$\langle x, y \rangle \equiv \int_{\theta \equiv (\theta_1, \dots, \theta_d) \in \mathcal{O}} x(\theta) y(\theta) d\theta, \quad x, y \in L^2(\mathcal{O}).$$

By a stochastic Allen–Cahn equation, we refer to the following formally written stochastic PDE

$$(1.8) \quad \begin{aligned} \frac{\partial}{\partial t} Y_n(t, \theta) &= \Delta Y_n(t, \theta) - V'(Y_n(t, \theta)) \\ &+ \frac{1}{\sqrt{n}} \sigma(\theta; Y_n(t)) \frac{\partial^{d+1}}{\partial t \partial \theta_1 \dots \partial \theta_d} \beta(t, \theta), \end{aligned}$$

where  $\beta(t, \theta)$  is a Brownian sheet over space–time  $(t, \theta) \in [0, \infty) \times \mathcal{O}$ ,  $V \in C^1(R)$  and

$$(1.9) \quad \sigma(\theta, y) = \varphi(\theta, \langle y, \xi_1 \rangle, \dots, \langle y, \xi_k \rangle)$$

for some  $\xi_1, \dots, \xi_k \in L^2(\mathcal{O})$  and  $\varphi(\theta, r_1, \dots, r_k): R^{k+1} \rightarrow R$ .

The above is a stochastically perturbed reaction–diffusion type equation. Among other applications, it has been used in material science as a phenomenological model of material interface movements due to molecular-level adsorption–desorption processes.  $V \in C^1(R)$  is usually a double- or multiple-well potential function. From large deviations for  $\{Y_n\}$ , we can extract information about metastability of the whole system when the temperature is small.

It is well known that, in dimension  $d \geq 2$ , (1.8) admits no  $L^2(\mathcal{O})$ -valued solution. We will actually consider an approximate version of it which is defined on truncated Fourier modes. The number of modes goes to infinity

as  $n \rightarrow +\infty$ . Such consideration is motivated by the fact that, in the above mentioned application, (1.8) should only be viewed as a formal limit for some stochastic Ginzburg–Landau equation defined on finite lattices or on truncated Fourier modes [30]. It is usually the rescaling limits of these finite systems which we really care about, rather than the continuum level (1.8). We mention that large deviation for the lattice case is studied in [18].

To rigorously define the processes, we let

$$\begin{aligned}\phi_1(r) &\equiv 1, & \phi_{2k-1}(r) &\equiv \sqrt{2} \cos(2\pi kr), \\ \phi_{2k}(r) &\equiv \sqrt{2} \sin(2\pi kr), & k &= 1, 2, \dots,\end{aligned}$$

and

$$\mu_1 \equiv 0, \quad \mu_{2k-1} = \mu_{2k} \equiv 4\pi^2 k^2.$$

It follows that

$$(1.10) \quad -\phi_j'' = \mu_j \phi_j.$$

Therefore

$$(1.11) \quad \begin{aligned} \{e_k(\theta) &\equiv \phi_{k_1}(\theta_1) \times \phi_{k_2}(\theta_2) \times \dots \times \phi_{k_d}(\theta_d), \\ k &\equiv (k_1, \dots, k_d), k_j = 1, 2, \dots; j = 1, \dots, d\} \end{aligned}$$

forms a complete orthonormal basis for  $E \equiv U_0 \equiv L^2(\mathcal{O})$ . Denote

$$(1.12) \quad \lambda_k \equiv \mu_{k_1} \times \dots \times \mu_{k_d};$$

then

$$-\Delta e_k = \lambda_k e_k.$$

Let  $\{\beta_k(t), k \equiv (k_1, \dots, k_d), k_j = 1, \dots, j = 1, \dots, d\}$  be a sequence of i.i.d. real-valued standard Brownian motion, and let

$$\beta(t, \theta) \equiv \sum_k \beta_k(t) \int_{r_1=0}^{\theta_1} \dots \int_{r_d=0}^{\theta_d} e_k(r_1, \dots, r_d) dr_1 \dots dr_d.$$

We define an  $L^2(\mathcal{O})$ -valued cylindrical Wiener process

$$(1.13) \quad W(t) \equiv \sum_k \beta_k(t) e_k.$$

Suppose

$$(1.14) \quad \lim_{n \rightarrow \infty} m_n = \infty, \quad \sup_n \frac{m_n^{4d}}{n} < \infty.$$

[This scaling is needed in (A.9) when verifying the exponential compact containment property (Condition 2.2) for the processes. In addition, it is also used to verify Condition 1.11(3) in Theorem 1.10.]



Let projection operator

$$P_n x \equiv \sum_{k_1=1}^{m_n} \cdots \sum_{k_d=1}^{m_n} \langle x, e_k \rangle e_k \in \text{span}(e_1, \dots, e_{m_n})$$

and for each  $x \in L^2(\mathcal{O})$  fixed, we define linear operator  $B(x)$  on  $L^2(\mathcal{O})$  by

$$(1.15) \quad (B(x)u)(\theta) \equiv \sigma(\theta; x)u(\theta), \quad u \in U_0 \equiv L^2(\mathcal{O}).$$

We regularize linear operator

$$(1.16) \quad B_n(x)u \equiv P_n(B(P_n x)u),$$

and arrive at an  $L^2(\mathcal{O})$ -valued diffusion

$$(1.17) \quad dX_n(t) = \Delta P_n X_n(t) dt - P_n V'(P_n X_n(t)) dt + \frac{1}{\sqrt{n}} B_n(X_n(t)) dW(t),$$

where the term  $B_n(X_n(t)) dW(t)$  is understood as

$$B_n(X_n(t)) dW(t) = \sigma(\cdot; P_n X_n(t)) \sum_{k_1=1}^{m_n} \cdots \sum_{k_d=1}^{m_n} d\beta_k(t) e_k.$$

Equation (1.17) can be written in the form of (1.7). Let  $\omega = \sup_{-\infty < r < \infty} |V''(r)|$  and

$$(1.18) \quad \begin{aligned} C_n x &= \Delta P_n x - P_n V'(P_n x) + P_n V'(0) - \omega P_n x, \quad x \in L^2(\mathcal{O}), \\ F_n(x) &= -P_n V'(0) + \omega P_n x. \end{aligned}$$

Then  $C_n 0 = 0$ ,  $C_n$  is  $m$ -dissipative in  $L^2(\mathcal{O})$  (Lemmas A.2 and A.3) and

$$dX_n(t) = C_n X_n(t) dt + F_n(X_n(t)) dt + \frac{1}{\sqrt{n}} B_n(X_n(t)) dW(t).$$

To prove the large deviation theorem, we assume:

CONDITION 1.3.

(1)  $V \in C^2(R)$  and

$$\sup_{-\infty < r < \infty} |V''(r)| < \infty.$$

(2) There exist  $c_1, c_2 > 0$ , such that

$$V(r) \geq c_1 + c_2 r^2.$$

(3) The  $\varphi(\theta, r_1, \dots, r_k): \mathcal{O} \times R^k \rightarrow R$  in (1.9) is bounded continuous and Lipschitz in  $r_1, \dots, r_k$ , uniformly with respect to  $\theta$ .

Condition 1.3(3) implies that operator  $B(x)$  defined by (1.15) is Lipschitz in  $x$ :

$$\sup_{x \neq y} \frac{\|B(x) - B(y)\|}{\|x - y\|_{L^2(\mathcal{O})}} < \infty,$$

where  $\|\cdot\|$  is the operator norm

$$\|B(x)\| \equiv \sup_{\|u\|_{L^2(\mathcal{O})} \leq 1} \|B(x)u\|_{L^2(\mathcal{O})}.$$

For each  $n$  fixed, (1.17) can actually be represented as a finite-dimensional stochastic ODE; therefore the existence and uniqueness of the solution hold by standard finite-dimensional results.

Applying the main theorem of this paper, Theorem 1.10, we have:

**THEOREM 1.4.** *Under Condition 1.3 and scaling relation (1.14), the solutions  $X_n(t)$  of (1.17) satisfy a large deviation principle in  $C_{L^2(\mathcal{O})}[0, \infty)$  with good rate function  $I$  as defined in (1.31).*

With a mild amount of additional work, assuming  $\inf_{\theta, x} \sigma(\theta, x) > 0$ , the rate function  $I$  can be represented more explicitly:

$$(1.19) \quad \begin{aligned} I(x) &= I_0(x(0)) \\ &+ \frac{1}{2} \int_0^\infty \int_{\mathcal{O}} \left| \frac{(\partial/\partial t)x(t, \theta) - \Delta_\theta x(t, \theta) + V'(x(t, \theta))}{\sigma(\theta, x(t))} \right|^2 d\theta dt. \end{aligned}$$

Feng and Kurtz [18] discuss this type of representation in general. See Section 4.1 for an outline of the approach.

**EXAMPLE 1.5** (Stochastic Cahn–Hilliard equation). We still consider  $\mathcal{O} = [0, 1]^d$  with periodic boundary condition, but with the restriction  $d = 1, 2$  or  $3$  now. We consider stochastic perturbation of the Cahn–Hilliard equation formally given by

$$(1.20) \quad \frac{\partial}{\partial t} Y_n(t, \theta) = \Delta(-\Delta Y_n(t, \theta) + V'(Y_n(t, \theta))) + \frac{1}{\sqrt{n}} \frac{\partial^{d+1}}{\partial t \partial \theta_1 \dots \partial \theta_d} \beta(t, \theta),$$

where  $\beta(t, \theta)$  is a Brownian sheet on  $[0, \infty) \times \mathcal{O}$ ; or, equivalently,

$$\frac{\partial}{\partial t} Y_n(t, \theta) + \operatorname{div}_\theta(\nabla(\Delta Y_n(t, \theta) - V'(Y_n(t, \theta)))) = \frac{1}{\sqrt{n}} \frac{\partial^{d+1}}{\partial t \partial \theta_1 \dots \partial \theta_d} \beta(t, \theta).$$

$Y_n$  is asymptotically conserved in the sense that, in the  $n \rightarrow +\infty$  limit  $Y$ ,  $\int Y(t, \theta) d\theta$  is constant in time. As in Example 1.2, such an equation has extensive applications in material science. Motivated by the same reason as before, we only rigorously study the large deviation for a variant of (1.20),

which is defined on finite Fourier modes. The number of modes goes to infinity slowly, as  $n$  goes to infinity.

We assume:

CONDITION 1.6.

(1)  $V \in C^3(R)$  and

$$\sup_{-\infty < r < \infty} |V''(r)| + |V'''(r)| < \infty.$$

(2) There exist  $c_1, c_2 > 0$ , such that

$$V(r) \geq c_1 + c_2 r^2.$$

(3)

$$(1.21) \quad \lim_{n \rightarrow \infty} m_n = \infty, \quad \sup_n \frac{m_n^{3d}}{n} < \infty.$$

The above scaling requirement on  $m_n$  is needed for reasons similar to those in the previous Allen–Cahn example. See the proof of (A.14) and the requirement in Condition 1.11(3).

We choose  $E = U_0 = L^2(\mathcal{O})$ . We define  $e_1, \dots, e_k, \dots$ ,  $P_n$  and  $W$  as in Example 1.2 and

$$B_n u \equiv P_n u \quad \forall u \in U_0 = L^2(\mathcal{O}).$$

We consider  $L^2(\mathcal{O})$ -valued diffusions:

$$(1.22) \quad dX_n(t) = \Delta P_n(-\Delta P_n X_n(t) dt + V'(P_n X_n(t))) dt + \frac{1}{\sqrt{n}} B_n dW(t),$$

where

$$B_n dW(t) = \sum_{k_1=1}^{m_n} \cdots \sum_{k_d=1}^{m_n} e_k d\beta_k(t).$$

Let  $\omega = \frac{1}{4} \sup_r |V''(r)|^2$ ,

$$(1.23) \quad C_n x \equiv \Delta P_n(-\Delta P_n x + V'(P_n x)) - \omega P_n x,$$

and

$$F_n(x) = \omega P_n x.$$

Then (1.22) can be written in the form of (1.7):

$$dX_n(t) = C_n X_n(t) dt + F_n(X_n(t)) dt + \frac{1}{\sqrt{n}} B_n dW(t).$$

THEOREM 1.7. *Under Condition 1.6 and the scaling relation (1.21), the solutions  $X_n(t)$  of (1.22) satisfy a large deviation principle in  $C_{L^2(\mathcal{O})}[0, \infty)$  with good rate function  $I$  as in (1.31).*

As in the stochastic Allen–Cahn example, the rate function  $I$  can be further simplified:

$$(1.24) \quad I(x) = I_0(x(0)) + \frac{1}{2} \int_0^\infty \int_{\mathcal{O}} \left| \frac{\partial}{\partial t} x(t, \theta) - \Delta_\theta(-\Delta_\theta x(t, \theta) + V'(x(t, \theta))) \right|^2 d\theta dt.$$

Another type of stochastic perturbation [30] to the Cahn–Hilliard equation could also be interesting:

$$\frac{\partial}{\partial t} Y_n(t, \theta) + \operatorname{div}_\theta \left( \nabla_\theta (\Delta_\theta Y_n(t, \theta) - V'(Y_n(t, \theta))) + \frac{1}{\sqrt{n}} \frac{\beta(\partial t, \partial \theta)}{\partial t \partial \theta} \right) = 0,$$

where

$$\beta(t, \theta) = (\beta_1(t, \theta), \dots, \beta_d(t, \theta))$$

with each  $\beta_k$  an independent real-valued space–time Brownian sheet. Large deviation for a lattice version of such an equation is considered in [18].

EXAMPLE 1.8 (Stochastic quasilinear equation with viscosity). Let  $\mathcal{O} = [0, 1)$  with periodic boundary. Suppose  $\phi \in C^1(R)$  and  $\sup_r |\phi'(r)| < \infty$ . We consider the following formally defined equation:

$$\frac{\partial}{\partial t} Y_n(t, \theta) + \frac{\partial}{\partial \theta} \phi(Y_n(t, \theta)) = \alpha \frac{\partial^2}{\partial \theta^2} Y_n(t, \theta) + \frac{1}{\sqrt{n}} \frac{\partial^2}{\partial t \partial \theta} \beta(t, \theta),$$

where  $\beta(t, \theta)$  is a Brownian sheet,  $(t, \theta) \in [0, \infty) \times \mathcal{O}$ .

As before, let  $E = U = L^2(\mathcal{O})$  with norm  $\|x\|^2 = \int_{\mathcal{O}} x^2(\theta) d\theta$ .  $\{e_1, \dots, e_k, \dots\}$  is the complete orthonormal basis as defined in (1.11) with  $d = 1$ . Define  $L^2(\mathcal{O})$ -valued cylindrical Wiener process

$$W(t) = \sum_{k=1}^{\infty} \beta_k(t) e_k$$

where  $\{\beta_1, \beta_2, \dots\}$  are i.i.d. standard Brownian motion. Let

$$(1.25) \quad \lim_{n \rightarrow \infty} m_n = \infty, \quad \sup_n \frac{m_n^3}{n} < \infty$$

and projection

$$P_n x = \sum_{k=1}^{2m_n} \langle x, e_k \rangle e_k.$$

We consider a regularized  $L^2(\mathcal{O})$ -valued diffusion equation,

$$(1.26) \quad dX_n(t) = \alpha \Delta_\theta P_n X_n(t) dt - P_n \partial_\theta \phi(P_n X_n(t)) dt + \frac{1}{\sqrt{n}} B_n dW(t),$$

where  $B_n = P_n$  and

$$B_n dW(t) = \sum_{k=1}^{2m_n} e_k d\beta_k(t).$$

The scaling on  $m_n$  is needed for the same reason as in the previous two examples.

Let  $\omega = (\sup_r \phi'(r))^2 / (4\alpha)$ ,

$$(1.27) \quad C_n x = \alpha \Delta_\theta P_n x - P_n \partial_\theta \phi(P_n x) - \omega P_n x, \quad x \in L^2(\mathcal{O}),$$

and

$$F_n(x) = \omega P_n x, \quad F(x) = \omega x.$$

Then  $X_n$  satisfies (1.7).

**THEOREM 1.9.** *Suppose that (1.25) holds and that  $\{X_n(0)\}$  satisfies a large deviation principle with rate function  $I_0$ . Then  $X_n \in C_{L^2(\mathcal{O})}[0, \infty)$  satisfy a large deviation principle with rate function given by (1.31).*

With additional work, it can be shown that the rate function  $I$  admits the following form:

$$I(x) = I_0(x(0)) + \frac{1}{2} \int_0^\infty \int_{\mathcal{O}} \left| \frac{\partial}{\partial t} x - \alpha \frac{\partial^2}{\partial \theta^2} x + \frac{\partial}{\partial \theta} \phi(x) \right|^2 d\theta dt.$$

**1.3. Technical assumptions and main results.** Let  $(E, \|\cdot\|)$  and  $(U_0, \|\cdot\|_{U_0})$  be two separable real Hilbert spaces. Let  $\{e_1, e_2, \dots\}$  be a complete orthonormal basis of  $U_0$ . We define  $W$ , a *cylindrical Wiener process* on  $U_0$ , by

$$(1.28) \quad W(t) = \sum_{k=1}^{\infty} e_k \beta_k(t), \quad t \geq 0,$$

where  $\beta_1, \beta_2, \dots$  are i.i.d. real-valued standard Brownian motions with respect to a filtration  $\mathcal{F}$ .  $X_n(0)$  is independent of  $\mathcal{F}$  and we write  $\mathcal{F}^n = \mathcal{F} \vee \sigma(X_n(0))$ . The infinite sum in (1.28) does not converge in  $(U_0, \|\cdot\|_{U_0})$ . However, we can always embed  $U_0$  continuously into another separable real Hilbert space  $(U_1, \|\cdot\|_{U_1})$ , and as far as the embedding is Hilbert–Schmidt, the right-hand side of (1.28) converges in  $(U_1, \|\cdot\|_{U_1})$ . For example, let  $\lambda_k > 0$  be such that  $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$ ; define  $U_1$  to be the completion of  $U_0$  under

$$\langle u, v \rangle_{U_1} = \sum_k \lambda_k^2 \langle v, e_k \rangle_{U_0} \langle u, e_k \rangle_{U_0}.$$

Then the embedding  $U_0 \rightarrow U_1$  through identity map  $J$  is Hilbert–Schmidt.

Throughout this paper, we will denote by  $L_2(U_0, E)$  the space of Hilbert–Schmidt operators from  $U_0$  into  $E$  with norm

$$(1.29) \quad \begin{aligned} \|B\|_{L_2(U_0, E)}^2 &\equiv \sum_k \|Be_k\|^2 \\ &= \text{Tr}(B^*B) = \text{Tr}(BB^*) \quad \forall B \in L_2(U_0, E). \end{aligned}$$

Therefore,  $J \in L_2(U_0, U_1)$ . To distinguish from this space, we will denote by  $L(U_0, E)$  the space of operators from  $U_0$  to  $E$  which are linear and bounded. For  $B \in L(U_0, E)$ ,

$$(1.30) \quad \|B\| = \|B\|_{L(U_0, E)} = \sup_{u \in U_0, \|u\| \leq 1} \|Bu\|.$$

The following stochastic integral will be used in this paper:

$$\int_0^t B(s) dW(s), \quad B(s) \in L_2(U_0, E).$$

Such an integral can be defined using telescoping Riemann summation just as the usual Itô integral in finite dimensions. Although the definition of  $W$  depends on  $U_1$ , the integral is independent of the choice of  $U_1$ . For details, see Chapter 4 of [9].

We identify an operator  $C$  in  $E$  by its graph:  $C \subset E \times E$ , and denote the space of continuous functions on  $E$  by  $C(E)$ .

Our main result in the paper is the following:

**THEOREM 1.10.** *Let  $E, U_0$  be arbitrary separable real Hilbert spaces. Suppose the following condition holds.*

**CONDITION 1.11.**

- (1) Operator  $C_n \subset E \times E$ ,  $F_n \in C(E)$  and  $B_n(x) \in L_2(U_0, E)$  are single valued and everywhere defined. That is,

$$\mathcal{D}(C_n) = \mathcal{D}(F_n) = E, \quad \mathcal{D}(B_n(x)) = U_0 \quad \forall x \in E.$$

Moreover, they are globally Lipschitz:

$$\begin{aligned} &\|C_n x - C_n y\| + \|F_n(x) - F_n(y)\| + \|B_n(x) - B_n(y)\|_{L_2(U_0, E)} \\ &\leq \text{Constant}_n \|x - y\|, \end{aligned}$$

for every  $x, y \in E$ , where the  $\text{Constant}_n$  may depend on  $n$ . [See (1.29) for the definition of  $\|\cdot\|_{L_2(U_0, E)}$ .]

- (2)  $C_n$  is  $m$ -dissipative on  $E$ ;  $C_n 0 = 0$  for  $n = 2, 3, \dots$ ; and there exists a (possibly multivalued)  $m$ -dissipative operator  $C \subset E \times E$ , with  $\overline{\mathcal{D}(C)} = E$  such that  $C \subset \lim_{n \rightarrow \infty} C_n$ , in the sense that for each  $(\xi, \eta) \in C$ , there exists  $\xi_n \in E$  such that  $\lim_n \|\xi - \xi_n\| + \|\eta - C_n \xi_n\| = 0$ .

(3) Whenever  $x_n \rightarrow x_0 \in E$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \| \| B_n(x_n) \| \|_{L_2(U_0, E)}^2 = 0.$$

(4) For each  $x \in E$ , there exist  $B(x) \in L(U_0, E)$  and  $F \in C(E)$  satisfying

$$\sup_{x \neq y} \frac{\| \| B(x) - B(y) \| \| + \| F(x) - F(y) \|}{\| x - y \|} < \infty,$$

where  $\| \| \cdot \| \|$  is the usual operator norm in (1.30). Furthermore, for each  $x_n, p_n \in E$  and  $x_n \rightarrow x_0, p_n \rightarrow p_0$ , we have

$$F_n(x_n) \rightarrow F(x_0) \quad \text{and} \quad \| B_n^*(x_n) p_n \|_{U_0} \rightarrow \| B^*(x_0) p_0 \|_{U_0}.$$

We also assume that  $X_n$  is the solution to (1.7), and that  $\{X_n : n = 1, 2, \dots\}$  satisfies the following exponential compact containment condition:

CONDITION 1.12. For each compact  $K \subset E$ ,  $T > 0$  and  $a > 0$ , there exists another compact set  $K_{a,T} \subset E$  such that

$$\limsup_{n \rightarrow \infty} \sup_{x \in K} \frac{1}{n} \log P(X_n(t) \notin K_{a,T}, \exists 0 < t \leq T | X_n(0) = x) \leq -a.$$

Finally, we assume that  $\{X_n(0) : n = 1, 2, \dots\}$  satisfies the large deviation principle with good rate function  $I_0$  on  $E$ .

Then:

- (a)  $\{X_n\}$  is exponentially tight;
- (b) the following limit exists and defines an operator semigroup on  $C_b(E)$  (the space of bounded continuous functions on  $E$ ):

$$V(t)f(x) = \lim_{n \rightarrow \infty, y \rightarrow x} \frac{1}{n} \log E[e^{nf(X_n(t))} | X_n(0) = y];$$

- (c) the large deviation principle holds for  $\{X_n\}$  with good rate function  $I$ :

$$(1.31) \quad I(x) = I_0(x(0)) + \sup_{0 \leq t_1 \leq \dots \leq t_m} \left( \sum_{i=1}^m I_{t_i - t_{i-1}}(x(t_i) | x(t_{i-1})) \right),$$

where

$$I_t(y|x) = \sup_{f \in C_b(E)} (f(y) - V(t)f(x)).$$

Theorems 1.4, 1.7 and 1.9 are all special cases of this theorem.

1.4. *Relation to other large deviation results in literature.* The term  $C_n$  in (1.7) and its limit  $C$  in Theorem 1.10 are allowed to be totally nonlinear. This is different than what is available in literature [4, 9, 24, 27, 29], where  $C$  is restricted to be semilinear. However, this paper does not pursue generalities in the term  $B_n$ , as some of the above mentioned papers do.

If  $C_n$  is semilinear, a good deal is known about the solution for (1.7). See [9]. However, if  $C_n$  is just  $m$ -dissipative, very little is known for the equation. By assuming  $C_n$  is Lipschitz and everywhere defined for each fixed  $n$ , we greatly simplified the situation. Such assumption is motivated by Examples 1.2 and 1.8, and by the fact that Yosida approximation of  $m$ -dissipative operators satisfies the above requirements.

In this paper (1.7) is driven by a Brownian noise  $W$ , which is responsible for the quadratic nonlinear term in  $H_0, H_1$  (or the  $\hat{H}_0, \hat{H}_1$ ). But in the proof of comparison principle, we actually allow much more general nonlinearity (Theorem 5.2). Therefore, it is possible that the method here can be applied to cases where the  $W$  is replaced by spatial Poisson noise. We expect exponential nonlinearity in the  $H_i, \hat{H}_i$ 's in these cases.

1.5. *Notation.* We will frequently use the following class of test functions for localization purpose:

$$(1.32) \quad \mathcal{T} = \{\varphi \in C^2([0, \infty)) : \varphi \geq 0, \text{ is nondecreasing,} \\ \text{and } \varphi(r) = \varphi(+\infty) \text{ for } r \text{ large enough}\}.$$

Throughout the paper,  $(E, r)$  and  $(U, r_U)$  are complete separable metric spaces. Let  $f$  be a function on  $E$ .  $C(E)$  denotes continuous functions on  $E$ ;  $C_b(E)$ , bounded continuous functions;  $B(E)$ , bounded Borel measurable functions;  $M(E)$ , Borel measurable functions; and  $\mathcal{P}(E)$ , probability measures on  $E$ . For  $f \in M(E)$ , we define  $f^*$  and  $f_*$  to be, respectively, the upper and lower semicontinuous smoothing of  $f$ :

$$f^*(x) = \limsup_{y \rightarrow x} f(y), \quad f_*(x) = \liminf_{y \rightarrow x} f(y).$$

Let  $\mathcal{O} \subset R^d$ :

$$H^1(\mathcal{O}) \equiv \left\{ x(\theta) \in L^2(\mathcal{O}) : \int_{\mathcal{O}} |x(\theta)|^2 + |\nabla x(\theta)|^2 d\theta < \infty \right\}$$

and

$$H^2(\mathcal{O}) \equiv \left\{ x(\theta) \in L^2(\mathcal{O}) : \sum_{i,j,k=1}^d \int_{\mathcal{O}} |x(\theta)|^2 + |\partial_k x(\theta)|^2 + |\partial_{i,j}^2 x(\theta)|^2 d\theta < \infty \right\}.$$

Throughout,  $(E, \|\cdot\|)$  is a real separable Hilbert space with its dual identified as itself  $E^* = E$ .  $C^k(E)$  denotes the set of  $k$ th-order Fréchet differentiable functions on  $E$  with continuous  $k$ th-order derivative; we identify



the  $k$ th-order derivative as a  $k$ th-order multilinear symmetric functional. For example, by  $Df(x)$ , we mean the gradient of  $f$  evaluated at  $x$ , which is identified as an element of  $E^* \equiv E$  through the Taylor expansion:

$$(1.33) \quad f(x+y) = f(x) + \langle Df(x), y \rangle + \frac{1}{2} D^2 f(x)yy + o(\|y\|^2), \quad y \in E.$$

$D^2 f(x)yz$  means a functional which is bilinear in both the  $y$  and the  $z$  arguments. Let  $x, y \in E$ ; by  $x \otimes y$ , we mean a bounded linear operator on  $E$

$$(x \otimes y)z \equiv x\langle y, z \rangle.$$

$|\cdot|$  will be used to denote either an absolute value of a number  $|a|$  or the Euclidean norm of a vector in  $R^d$ :  $|(\theta_1, \dots, \theta_d)|^2 = \sum_{k=1}^d \theta_k^2$ .

We denote the range of a generic operator  $A$  in a Banach space by  $\mathcal{R}(A)$  and its domain by  $\mathcal{D}(A)$ . We often identify an operator with its graph.  $\bar{A}$  denotes the closure of the graph under the norm of the Banach space. Let  $E$  be a Hilbert space. A possibly multivalued nonlinear operator  $C \subset E \times E$  is said to be *m-dissipative* if and only if it is *dissipative*:

$$\langle x_1 - x_2, y_1 - y_2 \rangle \leq 0 \quad \forall (x_i, y_i) \in C$$

and

$$\mathcal{R}(I - \alpha C) = E \quad \forall \alpha > 0.$$

If, in addition,  $\overline{\mathcal{D}(C)} = E$ , then  $C$  generates a strongly continuous contraction semigroup  $S(t)$  on  $E$ . The following test function on  $E$  is introduced to record trajectory properties of  $S(t)y$ . It plays a major role in the analysis of certain Hamilton–Jacobi equations (Section 5).

**DEFINITION 1.13.** Let  $C$  be an *m-dissipative* operator on  $E$  generating a strongly continuous semigroup  $S(t): t \geq 0$ . The *Tataru distance function*  $d_C$  is

$$d_C(x, y) \equiv \inf\{t + \|x - S(t)y\| : t \geq 0\} \quad \forall x, y \in E.$$

$d_C(x, y)$  is Lipschitz ((28) on page 62 of [8]):

$$|d_C(x, y) - d_C(\hat{x}, \hat{y})| \leq \|x - \hat{x}\| + \|y - \hat{y}\|.$$

Let  $E$  be a general metric space again and let  $H_0, H_1 \subset C_b(E) \times B(E)$  be (possibly multivalued) operators,  $h \in C_b(E)$  and  $\alpha > 0$ . We define viscosity solutions for

$$(1.34) \quad (I - \alpha H_0)f = h$$

and

$$(1.35) \quad (I - \alpha H_1)f = h.$$

DEFINITION 1.14 (Viscosity solution).

- (a)  $\bar{f}$  is a *viscosity subsolution* of (1.34) if and only if  $\bar{f}$  is bounded, upper semicontinuous, and for each  $(f_0, g_0) \in H_0$ , there exists an  $\{x_n\} \subset E$  satisfying

$$(1.36) \quad \lim_{n \rightarrow \infty} (\bar{f} - f_0)(x_n) = \sup_{x \in E} (\bar{f} - f_0)(x)$$

and

$$(1.37) \quad \limsup_{n \rightarrow \infty} (\alpha^{-1}(\bar{f} - h) - (g_0)^*)(x_n) \leq 0.$$

- (b)  $\underline{f}$  is a *viscosity supersolution* of (1.35) if and only if  $\underline{f}$  is bounded, lower semicontinuous, and for each  $(f_0, g_0) \in H_1$ , there exists an  $\{x_n\} \subset E$  satisfying

$$(1.38) \quad \lim_{n \rightarrow \infty} (f_0 - \underline{f})(x_n) = \sup_{x \in E} (f_0 - \underline{f})(x)$$

and

$$(1.39) \quad \liminf_{n \rightarrow \infty} (\alpha^{-1}(\underline{f} - h) - (g_0)_*)(y_n) \geq 0.$$

DEFINITION 1.15. We say a comparison principle holds for viscosity subsolution of (1.34) and supersolution of (1.35) if

$$\bar{f} \leq \underline{f},$$

for every subsolution  $\bar{f}$  of (1.34) and supersolution  $\underline{f}$  of (1.35).

Allowing  $H_0, H_1 \subset C(E) \times M(E)$ , Tataru [31, 32] and Crandall and Lions [8] define viscosity solution in a different manner. Definition 1.16 is an adaptation of their definitions when the domain of operator is chosen properly. We will explore the connection between such definition and the more general Definition 1.14, for equations arising in our large deviation context. See Section 3.4.

DEFINITION 1.16 (Tataru–Crandall–Lions).

- (a) We say that  $\bar{f}$  is a subsolution of (1.34), if  $\bar{f}$  is bounded upper semicontinuous on  $E$ , and for each  $x_0 \in E$  and  $f_0 \in \mathcal{D}(H_0)$  satisfying

$$(1.40) \quad (\bar{f} - f_0)(x_0) = \sup_{x \in E} (\bar{f} - f_0)(x),$$

we have

$$\alpha^{-1}(\bar{f} - h)(x_0) \leq (H_0 f_0)^*(x_0).$$

- (b) We say that  $\underline{f}$  is a supersolution of (1.35), if  $\underline{f}$  is bounded lower semicontinuous on  $E$ , and for each  $x_0 \in E$  and  $f_0 \in \mathcal{D}(H_1)$  satisfying

$$(1.41) \quad (f_0 - \underline{f})(x_0) = \sup_{x \in E} (f_0 - \underline{f})(x),$$

we have

$$\alpha^{-1}(\underline{f} - h)(x_0) \geq (H_1 f_0)_*(x_0).$$

DEFINITION 1.17 (Exponential tightness and large deviation principle). Let  $S$  be a complete separable metric space and let  $\{X_n\}$  be  $S$ -valued random variables.  $\{X_n\}$  is said to be *exponentially tight* if for every  $a > 0$ , there exists compact  $\mathcal{K}_a \subset S$  such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P(X_n \notin \mathcal{K}_a) < -a.$$

$\{X_n\}$  is said to satisfy the *large deviation principle* if there exists a lower semicontinuous function  $I: S \rightarrow [0, +\infty]$  such that for every open set  $A \subset S$ ,

$$-\inf_{x \in A} I(x) \leq \liminf_n \frac{1}{n} \log P(X_n \in A)$$

and for every closed set  $B \subset S$ ,

$$\limsup_n \frac{1}{n} \log P(X_n \in B) \leq -\inf_{x \in B} I(x).$$

$I$  is called the *rate function* and it is *good* if each level set is compact.

Let  $E$  be a complete separable metric space. For each  $n$ , let stochastic process  $X_n$  have state space  $E$  and let its trajectory be continuous in time. By large deviation (resp. exponential tightness) for the processes  $\{X_n\}$ , we apply the above definition with  $S = C_E[0, \infty)$ .

DEFINITION 1.18. Let  $(E, q)$  be a metric space.  $D \subset C_b(E)$  is said to *approximate the metric  $q$*  if for each compact  $K \subset E$  and  $z \in K$ , there exists  $f_n \in D$  such that  $\lim_{n \rightarrow \infty} \sup_{x \in K} |f_n(x) - q(x, z)| = 0$ .

**2. A general large deviation theorem.** This section presents a general theorem which is the basis for the large deviation method in this paper. The heuristics have been explained in Section 1.1.

Let  $(E, r)$  be a complete separable metric space and let  $\{X_n\}$  be a sequence of  $E$ -valued processes with trajectories in  $C_E[0, \infty)$ . Suppose  $A_n \subset B(E) \times B(E)$  is possibly multivalued. Let  $X_n$  be a solution to the  $A_n$ -martingale problem. That is, there is a filtration  $\mathcal{F}_t^n$ , such that

$$f(X_n(t)) - f(X_n(0)) - \int_0^t g(X_n(s)) ds \quad \forall (f, g) \in A_n$$

is a martingale. We will work under the following regularity condition.

CONDITION 2.1. For each  $n = 2, 3, \dots$ , let  $A_n \subset B(E) \times B(E)$ . We assume existence and uniqueness hold for the martingale problem for  $A_n$  with  $X_n \in C_E[0, +\infty)$  for each initial distribution  $\mu \in \mathcal{P}(E)$ . Let  $P_x^n \in \mathcal{P}(C_E[0, \infty))$  denote the distribution of the solution of the martingale problem for  $A_n$  with  $X_n(0) = x \in E$ ; we assume that the mapping  $x \rightarrow P_x^n$  is Borel measurable taking the weak topology on  $\mathcal{P}(C_E[0, \infty))$  (cf. Theorem 4.4.6 of [15]).

Define  $H_n \subset B(E_n) \times B(E_n)$  by

$$H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}, \quad e^{nf} \in \mathcal{D}(A_n),$$

or if  $A_n$  is multivalued,

$$H_n = \left\{ \left( f, \frac{1}{n} e^{-nf} g \right) : (e^{nf}, g) \in A_n \right\}.$$

The following is an exponential version of the *uniform compact containment* condition in the weak convergence theory.

CONDITION 2.2. For each compact  $K \subset E$ ,  $T > 0$  and  $a > 0$ , there exists a compact  $K_{a,T} \subset E$  such that

$$\limsup_{n \rightarrow \infty} \sup_{x \in K} \frac{1}{n} \log P(X_n(t) \notin K_{a,T}, \text{ for some } 0 \leq t \leq T | X_n(0) = x) \leq -a.$$

The following is an adaptation of Theorem 7.18 of [18]. In the adaptation, we also used a result of exponential tightness (Corollary 4.19), a variant of the Stone–Weierstrass theorem (Lemma A.8) and a technical estimate (Lemma 7.19). All the reference labels refer to [18].

THEOREM 2.3. *Let Condition 2.1 be satisfied. In addition, we assume the following:*

- (1) *Convergence of generators.* There exist  $H_0, H_1 \subset C_b(E) \times B(E)$  which are limits of the  $H_n$ 's in the following sense:
  - (a) For each  $(f, g) \in H_0$ , there exist some  $(f_n, g_n) \in H_n$  such that

$$\sup_n \left( \sup_x |f_n(x)| + \sup_x |g_n(x)| \right) < \infty,$$

and that for each  $x_n \rightarrow x_0$ , we have

$$\lim_{n \rightarrow \infty} f_n(x_n) = f(x_0), \quad \limsup_{n \rightarrow \infty} g_n(x_n) \leq g^*(x_0).$$

- (b) For each  $(f, g) \in H_1$ , there exist some  $(f_n, g_n) \in H_n$  [possibly different than those in (a)] such that

$$\sup_n \left( \sup_x |f_n(x)| + \sup_x |g_n(x)| \right) < \infty,$$

and that for each  $x_n \rightarrow x_0$ , we have

$$\lim_{n \rightarrow \infty} f_n(x_n) = f(x_0), \quad g_*(x_0) \leq \liminf_{n \rightarrow \infty} g_n(x_n).$$

There exist  $F \subset C_b(E)$  which approximate the metric  $q \equiv r \wedge 1$  (Definition 1.18), and for each  $f \in F$  and  $\lambda > 0$ ,  $\lambda f \in \mathcal{D}(H_0)$ .

- (2) Uniform exponential compact containment. Condition 2.2 holds.  
 (3) Comparison principle. There exist a subset  $D \subset C_b(E)$  and  $\alpha_0 > 0$ , such that for each  $h \in D$  and  $0 < \alpha < \alpha_0$ , the comparison principle (Definition 1.15) holds for subsolution (in the sense of Definition 1.14) of

$$(I - \alpha H_0)f = h,$$

and supersolution (Definition 1.14) of

$$(I - \alpha H_1)f = h.$$

$D$  contains an algebra that separates points and vanishes nowhere [i.e., for each  $x \in E$ , there exists  $f$  belonging to this algebra such that  $f(x) \neq 0$ ].

Define  $\{V_n(t)\}$  on  $B(E)$  by

$$V_n(t)f(x) = \frac{1}{n} \log E[e^{nf(X_n(t))} | X_n(0) = x].$$

If  $\{X_n(0)\}$  satisfies a large deviation principle in  $E$  with a good rate function  $I_0$ , then:

- (a) limit

$$V(t)f(x) \equiv \lim_{n \rightarrow \infty} V_n(t_n)f(x_n)$$

exists for every  $f \in C_b(E)$ ,  $t_n \rightarrow t$ , and  $x_n \rightarrow x$ .  $V(t)$  forms a nonlinear semigroup on  $C_b(E)$ .

- (b) for each  $0 \leq t_1 \leq \dots \leq t_k < \infty$ ,  $\{(X_n(t_1), \dots, X_n(t_k)) : n = 1, 2, \dots\}$  is exponentially tight in  $E^k$  and satisfies the large deviation principle with good rate function

$$\begin{aligned} & I_{t_1, \dots, t_k}(x_1, \dots, x_k) \\ &= \sup_{f_1, \dots, f_k \in D} \{f_1(x_1) + \dots + f_k(x_k) \\ & \quad - \Lambda_0(V(t_1))(f_1 + V(t_2 - t_1)) \end{aligned} \tag{2.1}$$

$$\begin{aligned} & \times (f_2 + \cdots + V(t_k - t_{k-1})f_k) \cdots) \} \\ & = \inf_{x_0 \in E} \left\{ I_0(x_0) + \sum_{i=1}^k I_{t_i - t_{i-1}}(x_i | x_{i-1}) \right\}, \end{aligned}$$

where

$$\Lambda_0(f) = \lim_{n \rightarrow +\infty} \frac{1}{n} \log E[e^{nf(X_n(0))}] \quad \forall f \in C_b(E)$$

[the limit exists by (1.1)], and

$$(2.2) \quad I_t(y|x) = \sup_{f \in C_b(E)} (f(y) - V(t)f(x)).$$

(c)  $\{X_n\}$  is exponentially tight in  $C_E[0, \infty)$  and satisfies the large deviation principle with good rate function:

$$(2.3) \quad \begin{aligned} I(x) &= \sup_{k=1,2,\dots} I_{t_1,\dots,t_k}(x(t_1), \dots, x(t_k)) \\ &= \sup_{k=1,2,\dots} \sup_{0 < t_1 < t_2 < \dots < t_k} \left( I_0(x(0)) + \sum_{i=1}^k I_{t_i - t_{i-1}}(x(t_i) | x(t_{i-1})) \right). \end{aligned}$$

REMARK 2.4. In view of the duality in (1.2), the form of rate function in the first identity of (2.1) should be expected. The second equality follows by the Markovian property of the  $X_n$ 's. The rate function (2.3) follows from the finite-dimensional large deviation result in (b), and from a well-known projective limit argument [10, 11, 28].

Large deviation behavior of the  $\{X_n\}$  and the exponential tightness imply that the rate function  $I$  is good. That is,  $I$  has compact level sets. See part (b) of Lemma 1.2.18 of [12]. Similarly,  $I_{t_1,\dots,t_k}$  has compact level sets in  $E^k$ .

Theorem 2.3 can be applied to situations other than small perturbation type problems; we refer the reader to [18] for further examples.

In the rest of this paper we apply the above theorem to the general problem considered in Theorem 1.10. Step 1 is verified in Section 3.3; see Lemma 3.4. The condition in step 2 is assumed in Theorem 1.10 as Condition 1.12. It is verified for Examples 1.2, 1.5 and 1.8 in Section A.2. The comparison principle in step 3 is stated in Lemma 3.10, with details of the actual proof carried out in Section 5.

Before closing this section, we illustrate how the classical Freidlin–Wentzell theory follows from Theorem 2.3.

EXAMPLE 2.5 (The Freidlin–Wentzell theory). Let  $E = R^d$ , and let  $W$  be a standard  $d$ -dimensional Brownian motion. Assume that  $b_j, \sigma_{ij} \in C_b(R^d)$

are Lipschitz continuous for  $i, j = 1, \dots, d$ . We denote  $b(x) = (b_1(x), \dots, b_d(x)) : R^d \rightarrow R^d$ , and let  $d \times d$ -matrix  $\sigma(x) = (\sigma_{ij}(x))$ . Let  $X_n$  be the solution to

$$dX_n(t) = b(X_n(t)) dt + \frac{1}{\sqrt{n}} \sigma(X_n(t)) dW(t).$$

This is a special case of (1.7).

Let  $D = \{f : f = f_0 + c, f_0 \in C_0^2(R^d), c \in R\}$  where  $C_0^2(R^d)$  is the collection of functions with compact support and with continuous derivative up to the second order. We denote by  $D^2 f(x) = (\partial_{ij}^2 f(x))_{i,j}$  the Hessian matrix of  $f$ . By Itô's formula, if we take

$$A_n f(x) = b(x) \nabla f(x) + \frac{1}{2n} \text{Tr}(D^2 f(x) \sigma(x) \sigma^T(x)), \quad f \in D,$$

then Condition 2.1 is satisfied. The transformed generator

$$H_n f(x) = b(x) \nabla f(x) + \frac{1}{2} |\sigma^T(x) \nabla f(x)|^2 + \frac{1}{2n} \text{Tr}(D^2 f(x) \sigma(x) \sigma^T(x)),$$

$f \in D.$

If we let  $H_0 = H_1 = H$  with

$$Hf(x) = b(x) \nabla f(x) + \frac{1}{2} |\sigma^T(x) \nabla f(x)|^2, \quad f \in D,$$

then the convergence conditions in part (1) of Theorem 2.3 are satisfied.

The assumptions on  $\sigma, b$  imply that they grow at most linearly:

$$\sum_{i,j,k=1}^d (|\sigma_{ij}(x)| + |b_k(x)|) \leq c_1 + c_2 |x|.$$

One can use such estimate to verify the uniform exponential compact containment condition in Theorem 2.3. This is shown in Example 4.23 of [18] using a stochastic Lyapunov function technique.

Let  $h \in C_b(R^d)$  and  $\alpha > 0$ ; the comparison principle for

$$(2.4) \quad (I - \alpha H)f = h$$

follows from results in [5]. Details on its proof can also be found in Chapters 9.4 and 10.3 of [18].

Consequently, by Theorem 2.3, the large deviation principle holds for  $\{X_n\}$ .

Let  $\int_0^T |u(s)|^2 ds < +\infty$  for each  $T > 0$  and consider

$$(2.5) \quad \dot{x}(t) = b(x(t)) + \sigma(x(t))u(t).$$

We define

$$R_\alpha h(x_0) = \sup \left\{ \int_0^\infty e^{-\alpha^{-1}s} (\alpha^{-1} h(x(s)) - \frac{1}{2} |u(s)|^2) ds : (x, u) \text{ satisfies (2.5), } x(0) = x_0 \right\}.$$

Then it can be shown that  $R_\alpha h \in C_b(R^d)$ , it is the unique solution to (2.4), and

$$\begin{aligned} V(t)h(x_0) &= \lim_{k \rightarrow +\infty} R_{t/k}^k h(x_0) \\ &= \sup \left\{ \int_0^t \frac{1}{2} |u(s)|^2 ds + h(x(t)) : \dot{x} = b(x) + \sigma(x)u, x(0) = x_0 \right\}. \end{aligned}$$

All these can be rigorously justified using the dynamic programming principle. Suppose that  $\sigma^{-1}(x)$  exists. Plug the above expression on  $V$  in (2.3) and (2.2); we obtain the simplified representation

$$I(x) = \int_0^\infty |\sigma^{-1}(x)(\dot{x} - b(x))|^2 ds.$$

Such result is known as the Freidlin–Wentzell theory [23].

All the above claims are well-known results in control theory and first-order Hamilton–Jacobi equation literature. In [18], rigorous proofs are provided and summarized again.

**3. Large deviation for diffusions in Hilbert space.** Throughout this section we assume Condition 1.11 holds and  $X_n$  is a solution of (1.7). Both  $E$  and  $U_0$  are real separable Hilbert spaces.

**3.1. Semigroup on Hilbert spaces.** We first recall some basic facts about semigroup on  $E$ . These facts will be used later in the paper.

We assumed, in Condition 1.11(2), that  $C$  is  $m$ -dissipative on  $E$ , and  $\overline{\mathcal{D}(C)} = E$ . By Crandall and Liggett's [6] semigroup generation theorem, it generates a strongly continuous contraction semigroup on  $E$ :

$$S(t)x = \lim_{n \rightarrow \infty} \left( I - \frac{t}{n}C \right)^{-n} x \quad \forall x \in E,$$

and

$$\|S(t)x - S(t)y\| \leq \|x - y\| \quad \forall t > 0, x, y \in E.$$

Since  $(0, 0) \in C$  (or  $C0 = 0$  if  $C$  is single valued),  $0 = (I - \alpha C)^{-1}0$ ,  $S(t)0 = 0$  and  $\|S(t)x\| \leq \|x\|$ .

**DEFINITION 3.1** (Canonical restriction of  $C$ ). We denote

$$\|Cx\| \equiv \inf\{\|y\| : (x, y) \in C\} \quad \forall x \in \mathcal{D}(C)$$

and define a single-valued  $C^0 \subset E \times E$ , called *the canonical restriction of  $C$* , by

$$C^0x = \{z : (x, z) \in C, \|z\| = \|Cx\|\}.$$



Then the following holds.

LEMMA 3.2.

- (1)  $\mathcal{D}(C^0) = \mathcal{D}(C)$  and  $C^0$  is single valued (Lemma 2.19 of [26]).
- (2)  $C^0$  is the infinitesimal generator of  $S(t)$  in the sense that

$$(3.1) \quad C^0 x = \lim_{h \rightarrow 0+} \frac{1}{h} (S(h)x - x), \quad x \in \mathcal{D}(C)$$

(Corollary 4.19 in [26]).

- (3) Let  $f \in C^1(E)$ ; then

$$(3.2) \quad \langle Df(\xi), C^0 \xi \rangle = \lim_{r \rightarrow 0+} \frac{f(S(r)\xi) - f(\xi)}{r} \quad \forall \xi \in \mathcal{D}(C).$$

DEFINITION 3.3 [Directional derivative along the trajectory of  $S(t)$ ]. Suppose  $f \in C(E)$  is Lipschitz continuous. We define

$$D_C^+ f(x) = \limsup_{h \rightarrow 0+, y \rightarrow x} \frac{1}{h} (f(S(h)y) - f(y))$$

and

$$D_C^- f(x) = \liminf_{h \rightarrow 0+, y \rightarrow x} \frac{1}{h} (f(S(h)y) - f(y)).$$

$D_C^+ f : E \rightarrow [-\infty, +\infty]$  is upper semicontinuous, and  $D_C^- f : E \rightarrow [-\infty, +\infty]$  is lower semicontinuous (Lemma 2.3 in [8]).

We list two useful properties of the Tataru distance function  $d_C$  (Definition 1.13):

$$(3.3) \quad d_C(x, y) - d_C(\hat{x}, \hat{y}) \leq \|x - \hat{x}\| + \|y - \hat{y}\|$$

and

$$(3.4) \quad \frac{d_C(S(r)x, y) - d_C(x, y)}{r} \leq 1, \quad D_C^+ d_C(\cdot, y) \leq 1.$$

See page 62 of [8] for proof.

3.2. *The martingale problem.* Recall that  $E$  is a real separable Hilbert space, and that  $C_n$  is single valued, everywhere defined and Lipschitz on the  $E$  ( $C_n$  is usually some regularization of the  $C$ ). Let  $f \in C^2(E)$  be such that  $Df(x) = 0$  when  $\|x\|$  is sufficiently large; we define linear operator  $A_n \subset C_b(E) \times B(E)$  by

$$(3.5) \quad A_n f(x) = \langle Df(x), C_n x + F_n(x) \rangle + \frac{1}{2n} \text{Tr}[D^2 f(x) B_n(x) B_n^*(x)].$$

By Condition 1.11, an infinite-dimensional version of the Itô formula applies (e.g., Theorem 4.17 of [9]). In addition, Condition 1.11(1) is a strong enough assumption so that the results in Chapter 9 of [9] (regarding Markov property and regularity for the initial conditions) apply. Therefore Condition 2.1 is satisfied.

We next compute nonlinear operator  $H_n \subset C_b(E) \times B(E)$  by

$$\begin{aligned}
 H_n f(x) &\equiv \frac{1}{n} e^{-nf} A_n e^{nf}(x) \\
 &= \frac{1}{n} e^{-nf}(x) \langle D e^{nf}(x), C_n x + F_n(x) \rangle \\
 &\quad + \frac{1}{2n^2} e^{-nf(x)} \text{Tr}[D^2 e^{nf}(x) B_n(x) B_n^*(x)] \\
 (3.6) \quad &= \langle Df(x), C_n x + F_n(x) \rangle \\
 &\quad + \frac{1}{2n^2} \text{Tr}[(D(nf)(x) \otimes D(nf)(x) + D^2(nf)(x)) B_n(x) B_n^*(x)] \\
 &= \langle Df(x), C_n x + F_n(x) \rangle \\
 &\quad + \frac{1}{2} \|B_n^*(x) Df(x)\|_{U_0}^2 + \frac{1}{2n} \text{Tr}[D^2 f(x) B_n(x) B_n^*(x)],
 \end{aligned}$$

where  $(x \otimes y)z \equiv x \langle y, z \rangle$ . The last step above needs some justification: let  $\{\hat{e}_1, \dots, \hat{e}_k, \dots\}$  be a complete orthonormal basis of  $E$ . Then

$$\begin{aligned}
 &\frac{1}{2n^2} \text{Tr}[(D(nf)(x) \otimes D(nf)(x)) B_n(x) B_n^*(x)] \\
 &= \frac{1}{2} \sum_{k=1}^{\infty} \langle (Df(x) \otimes Df(x)) B_n(x) B_n^*(x) \hat{e}_k, \hat{e}_k \rangle \\
 &= \frac{1}{2} \sum_{k=1}^{\infty} \langle Df(x), \hat{e}_k \rangle \langle Df(x), B_n(x) B_n^*(x) \hat{e}_k \rangle \\
 &= \frac{1}{2} \langle Df(x), B_n(x) B_n^*(x) Df(x) \rangle \\
 &= \frac{1}{2} \|B_n^*(x) Df(x)\|_{U_0}^2.
 \end{aligned}$$

**3.3. Convergence of the  $H_n$ 's.** Formally, we expect the limit of  $H_n f$  to be given by

$$Hf(x) = \langle Df(x), Cx + F(x) \rangle + \frac{1}{2} \|B^*(x) Df(x)\|_{U_0}^2.$$

However, the above does not make sense for  $x \notin \mathcal{D}(C)$ . As commented in Section 1.1, we have to replace  $H$  by  $H_0, H_1$ ; then by selecting test functions  $f$  carefully, we can estimate the limit from above by  $H_0 f$  and from below

by  $H_1 f$ . The class of test functions has to be large enough so that the comparison principle (Sections 3.4 and 5) can be proved.

This is what we will carry out rigorously next.

By Condition 1.11,  $C_n$  and  $C$  generate, respectively, strongly continuous contraction semigroup  $S_n(t)$  and  $S(t)$  on  $E$ ,  $S_n(t)0 = 0$ ,  $S(t)0 = 0$  and  $\|S_n(t)x\| \leq \|x\|$ ,  $\|S(t)x\| \leq \|x\|$ .

We derive the limit operators  $H_0, H_1$  in Theorem 2.3 through several steps.

First, recall definitions of the *canonical restriction of  $C$*  in Definition 3.1 and of the *Tataru distance function  $d_C$*  in Definition 1.13.  $d_C$  is Lipschitz; however, it may not be differentiable in  $x$ . We introduce smooth approximations of it first.

By Condition 1.11, both  $C_n$  and  $C$  are  $m$ -dissipative, and  $C \subset \lim_n C_n$ . By the Crandall–Liggett semigroup convergence theorem ([6]; see also Theorem 6.8 of [26]),

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} \|S_n(t)y - S(t)y\| = 0 \quad \forall y \in E, T > 0.$$

Let  $\lim_{n \rightarrow \infty} a_n = \infty$ ; we define

$$(3.7) \quad \phi_\varepsilon(r) = \left( \sqrt{\varepsilon} + \frac{r - \varepsilon}{2\sqrt{\varepsilon}} - \frac{(r - \varepsilon)^2}{8\varepsilon\sqrt{\varepsilon}} \right) I(0 \leq r < \varepsilon) + \sqrt{r} I(r \geq \varepsilon),$$

$$(3.8) \quad h_{\varepsilon,y}(x) \equiv \inf_{t \geq 0} \{t + \phi_\varepsilon(\|x - S(t)y\|^2)\},$$

$$(3.9) \quad h_{n,\varepsilon,y}(x) \equiv -\frac{1}{a_n} \log \int_0^\infty e^{-a_n \{t + \phi_\varepsilon(\|x - S_n(t)y\|^2)\}} dt.$$

Then by Lemma A.12,

$$\lim_{\varepsilon \rightarrow 0+} \sup_{x \in E} |h_{\varepsilon,y}(x) - d_C(x, y)| = 0$$

and

$$\lim_{n \rightarrow \infty} \sup_{x \in K} |h_{\varepsilon,y}(x) - h_{n,\varepsilon,y}(x)| = 0$$

for each compact  $K \subset E$ . Later, we may drop the  $y$  in the subindex if no confusion can occur.

Recall the definition of  $\mathcal{T}$  in (1.32). We now define  $H_0$  and  $H_1$ :

(a) Let

$$\begin{aligned} \mathcal{D}(H_0) = \{f(x) : f(x) = \varphi_1(\|x - \xi\|^2) + \varphi_2(h_{\varepsilon,y_1}(x)) + \cdots + \varphi_{k+1}(h_{\varepsilon,y_k}(x)), \\ \forall \varphi_i \in \mathcal{T}, \xi \in \mathcal{D}(C), y_j \in E, k = 1, 2, \dots\}. \end{aligned}$$

For  $g(x) = \varphi_1(\|x - \xi\|^2)$  and  $f(x) = g(x) + \varphi_2(h_{\varepsilon, y_1}(x)) + \cdots + \varphi_{k+1}(h_{\varepsilon, y_k}(x)) \in \mathcal{D}(H_0)$ , we define

$$\begin{aligned}
 H_0 f(x) &= 2\varphi'_1(\|x - \xi\|^2)\langle x - \xi, C^0 \xi \rangle \\
 &\quad + \left( \sup_{r \geq 0} \varphi'_2(r) + \cdots + \sup_{r \geq 0} \varphi'_{k+1}(r) \right) \\
 (3.10) \quad &\quad + \sup_{\|q\| \leq \varphi'_2(h_{\varepsilon, y_1}(x)) + \cdots + \varphi'_{k+1}(h_{\varepsilon, y_k}(x))} (\langle F(x), Dg(x) + q \rangle \\
 &\quad + \frac{1}{2} \|B^*(x)(Dg(x) + q)\|_{U_0}^2).
 \end{aligned}$$

By item (1) of Lemma A.12, and the fact that  $Dg(x) = 0$  when  $\|x\|$  is sufficiently large, we have

$$\begin{aligned}
 &\sup_{x \in E} \sup_{\|q\| \leq \varphi'_2(h_{\varepsilon, y_1}(x)) + \cdots + \varphi'_{k+1}(h_{\varepsilon, y_k}(x))} \|\langle F(x), Dg(x) + q \rangle\| \\
 &\quad + \frac{1}{2} \|B^*(x)(Dg(x) + q)\|_{U_0}^2 < \infty.
 \end{aligned}$$

Consequently,  $H_0 \subset C_b(E) \times B(E)$ .

(b) Let

$$\begin{aligned}
 \mathcal{D}(H_1) &= \{f(x) : f(x) = -\varphi_1(\|x - \xi\|^2) \\
 &\quad - \varphi_2(h_{\varepsilon, y_1}(x)) - \cdots - \varphi_{k+1}(h_{\varepsilon, y_k}(x)), \\
 &\quad \forall \varphi_i \in \mathcal{T}, \xi \in \mathcal{D}(C), y_j \in E, k = 1, 2, \dots\}.
 \end{aligned}$$

Let  $f(x) = g(x) - \varphi_2(h_{\varepsilon, y_1}(x)) - \cdots - \varphi_{k+1}(h_{\varepsilon, y_k}(x)) \in \mathcal{D}(H_1)$ , where  $g(x) = -\varphi_1(\|x - \xi\|^2)$ . We define

$$\begin{aligned}
 H_1 f(x) &= -2\varphi'_1(\|x - \xi\|^2)\langle x - \xi, C^0 \xi \rangle \\
 &\quad - \left( \sup_{r \geq 0} \varphi'_2(r) + \cdots + \sup_{r \geq 0} \varphi'_{k+1}(r) \right) \\
 (3.11) \quad &\quad + \inf_{\|q\| \leq \varphi'_2(h_{\varepsilon, y_1}(x)) + \cdots + \varphi'_{k+1}(h_{\varepsilon, y_k}(x))} (\langle F(x), Dg(x) + q \rangle \\
 &\quad + \frac{1}{2} \|B^*(x)(Dg(x) + q)\|_{U_0}^2).
 \end{aligned}$$

The reason for using  $\varphi_i$  is to localize the test function  $f$  and the  $H_0 f, H_1 f$  so that  $H_0, H_1 \subset C_b(E) \times B(E)$ , a condition required by Theorem 2.3. This is the main reason that the two operators have such complicated forms, instead of the simpler forms used by Crandall and Lions [8] using the  $D_C^+$  in Definition 3.3. We note that  $H_0$  and  $H_1$  are both single valued.

Let

$$(3.12) \quad F = \{f(x) = \varphi_1(\|x - \xi\|^2), \varphi \in \mathcal{T}, \xi \in \mathcal{D}(C)\} \subset \mathcal{D}(H_0).$$

Then  $F$  approximates the metric  $q(x, y) = \|x - y\| \wedge 1$ . In addition, for  $\lambda > 0$ , if  $f \in \mathcal{D}(H_0)$ , then  $\lambda f \in \mathcal{D}(H_0)$ .

LEMMA 3.4.

- (1) For each  $f \in \mathcal{D}(H_0)$ , there exists  $f_n \in \mathcal{D}(H_n)$  such that

$$\sup_n \sup_x (|f_n(x)| + |H_n f_n(x)|) < \infty$$

and

$$\lim_{n \rightarrow +\infty} f_n(x_n) = f(x_0),$$

$$\limsup_{n \rightarrow +\infty} H_n f_n(x_n) \leq (H_0 f)^*(x_0)$$

whenever  $x_n \rightarrow x_0$ .

- (2) For each  $f \in \mathcal{D}(H_1)$ , there exists  $f_n \in \mathcal{D}(H_n)$  such that

$$\sup_n \sup_x (|f_n(x)| + |H_n f_n(x)|) < \infty$$

and

$$\lim_{n \rightarrow +\infty} f_n(x_n) = f(x_0),$$

$$\liminf_{n \rightarrow +\infty} H_n f_n(x_n) \geq (H_1 f)_*(x_0)$$

whenever  $x_n \rightarrow x_0$ .

PROOF. Let us present the proof for  $H_0$  only; the case for  $H_1$  is similar. To further simplify, let us just verify the case for test functions in  $\mathcal{D}(H_0)$  of the form

$$f(x) = g(x) + \varphi_2(h_{\varepsilon,y}(x)) \equiv \varphi_1(\|x - \xi\|^2) + \varphi_2(h_{\varepsilon,y}(x)) \in \mathcal{D}(H_0),$$

where  $h_{\varepsilon,y}$  is defined by (3.8). By Condition 1.11, there exists  $\xi_n \in E$  such that

$$\lim_{n \rightarrow \infty} \|\xi - \xi_n\| + \|C^0 \xi - C_n \xi_n\| = 0.$$

Let  $a_n > 1$  satisfy  $\lim_{n \rightarrow \infty} a_n = +\infty$ . We define  $h_{n,\varepsilon,y}$  according to (3.9). Let

$$g_n(x) \equiv \varphi_1(\|x - \xi_n\|^2),$$

$$f_n(x) \equiv g_n(x) + \varphi_2(h_{n,\varepsilon,y}(x)) \in \mathcal{D}(H_n).$$

By part 3 of Lemma A.12,

$$\lim_{n \rightarrow \infty} \sup_{x \in K} |f_n(x) - f(x)| = 0 \quad \text{for each } K \subset E \text{ compact.}$$

Apply (3.2), (A.18) and (A.20) to (3.6):

$$\begin{aligned}
(3.13) \quad H_n f_n(x) &\leq 2\varphi'_1(\|x - \xi_n\|^2) \langle x - \xi_n, C_n \xi_n \rangle \\
&\quad + \sup_{r \geq 0} \varphi'_2(r) + \langle F_n(x), D(g_n + \varphi_2 \circ h_{n,\varepsilon,y})(x) \rangle \\
&\quad + \frac{1}{2} \|B_n^*(x) D(g_n + \varphi_2 \circ h_{n,\varepsilon,y})(x)\|_{U_0}^2 \\
&\quad + \frac{1}{2n} \operatorname{Tr}[D^2 g_n(x) B_n(x) B_n^*(x)] \\
&\quad + \frac{1}{2n} \operatorname{Tr}[D^2(\varphi_2 \circ h_{n,\varepsilon,y_1})(x) B_n(x) B_n^*(x)] \\
&\leq 2\varphi'_1(\|x - \xi_n\|^2) \langle x - \xi_n, C_n \xi_n \rangle + \sup_{r \geq 0} \varphi'_2(r) \\
&\quad + \sup_{\|q\| \leq \varphi'_2(h_{n,\varepsilon,y}(x))} \left( \langle F_n(x), Dg_n(x) + q \rangle \right. \\
&\quad \quad \quad \left. + \frac{1}{2} \|B_n^*(x) (Dg_n(x) + q)\|_{U_0}^2 \right) \\
&\quad + \frac{1}{2n} \operatorname{Tr}[D^2 g_n(x) B_n(x) B_n^*(x)] \\
&\quad + \frac{1}{2n} \operatorname{Tr}[D^2(\varphi_2 \circ h_{n,\varepsilon,y})(x) B_n(x) B_n^*(x)],
\end{aligned}$$

where

$$(3.14) \quad Dg_n(x) = 2\varphi'_1(\|x - \xi_n\|^2)(x - \xi_n),$$

$$(3.15) \quad D^2 g_n(x) = 2\varphi'_1(\|x - \xi_n\|^2)I + 2\varphi''_1(\|x - \xi_n\|^2)(x - \xi_n) \otimes (x - \xi_n).$$

Let  $x_n \rightarrow x_0$ , and denote

$$\delta_n = \frac{1}{n} \|B_n(x_n)\|_{L_2(U_0, E)}^2.$$

Then  $\delta_n \rightarrow 0$  according to Condition 1.11(3). Taking  $a_n = \delta_n^{-1/2}$  to be the one in (3.9), then  $a_n \rightarrow +\infty$  and

$$\frac{1}{n} a_n \|B_n(x_n)\|_{L_2(U_0, E)}^2 = a_n \delta_n = \delta_n^{1/2} \rightarrow 0.$$

By (A.21),

$$(3.16) \quad \lim_{n \rightarrow \infty} \frac{1}{2n} \operatorname{Tr}[D^2(\varphi_2 \circ h_{n,\varepsilon,y})(x_n) B_n(x_n) B_n^*(x_n)] = 0.$$

Therefore, by (3.13) through (3.16),

$$\limsup_{n \rightarrow +\infty} H_n f_n(x_n) \leq (H_0 f)^*(x_0)$$

whenever  $x_n \rightarrow x_0 \in E$ .  $\square$

3.4. *The comparison principle.* Let  $\alpha > 0$ , and let  $h \in C_b(E)$  be uniformly continuous on  $E$ . The main goal of this subsection is to prove the comparison principle in Lemma 3.10.

In what follows, we extend the operator  $H_0, H_1$  and connect Feng and Kurtz's definition of viscosity solution (Definition 1.14) with those in [31, 32] and [8]. We will introduce a new set of operators  $\tilde{H}_0, \tilde{H}_1, \hat{H}_0, \hat{H}_1$  and will denote  $\bar{H}_0, \bar{H}_1$  closures of  $H_0, H_1$  under the graph norm topology in  $B(E)$ . We will clarify the relations among the next four sets of equations:

$$(3.17) \quad (I - \alpha H_0)f = h,$$

$$(3.18) \quad (I - \alpha H_1)f = h;$$

$$(3.19) \quad (I - \alpha \bar{H}_0)f = h,$$

$$(3.20) \quad (I - \alpha \bar{H}_1)f = h;$$

$$(3.21) \quad (I - \alpha \tilde{H}_0)f = h,$$

$$(3.22) \quad (I - \alpha \tilde{H}_1)f = h$$

and

$$(3.23) \quad (I - \alpha \hat{H}_0)f = h,$$

$$(3.24) \quad (I - \alpha \hat{H}_1)f = h.$$

Let  $\varphi_i \in \mathcal{T}$  [see (1.32)],  $y_i \in E$  and  $\xi \in \mathcal{D}(C)$ ,

$$g(x) = \varphi_1(\|x - \xi\|^2),$$

$$f(x) = g(x) + \varphi_2(d_C(x, y_1)) + \cdots + \varphi_{k+1}(d_C(x, y_k));$$

we define single-valued operator

$$(3.25) \quad \begin{aligned} \tilde{H}_0 f(x) &= 2\varphi'_1(\|x - \xi\|^2)\langle x - \xi, C^0 \xi \rangle \\ &\quad + \left( \sup_{r \geq 0} \varphi'_2(r) + \cdots + \sup_{r \geq 0} \varphi'_{k+1}(r) \right) \\ &\quad + \sup_{\|q\| \leq \varphi'_2(d_C(x, y_1)) + \cdots + \varphi'_{k+1}(d_C(x, y_k))} (\langle F(x), Dg(x) + q \rangle \\ &\quad + \tfrac{1}{2} \|B^*(x)(Dg(x) + q)\|_{\mathcal{U}_0}^2). \end{aligned}$$

By item (1) in Lemma A.12, (3.10) is equal to (3.25) when  $\|x\|$  is sufficiently large, independent of the  $\varepsilon$ . In addition, by (A.17),

$$\lim_{\varepsilon \rightarrow 0+} \sup_{x \in E} |h_{\varepsilon, y}(x) - d_C(x, y)| = 0.$$

Therefore sending  $\varepsilon \rightarrow 0$ , we obtain  $\tilde{H}_0 \subset \bar{H}_0$ , where  $\bar{H}_0$  is the closure of  $H_0$  under the uniform norm for  $B(E)$ . Similarly, let

$$\begin{aligned} g(x) &= -\varphi_1(\|x - \xi\|^2), \\ f(x) &= g(x) - (\varphi_2(d_C(x, y_1)) + \cdots + \varphi_{k+1}(d_C(x, y_k))) \end{aligned}$$

and define

$$\begin{aligned} \tilde{H}_1 f(x) &= -2\varphi'_1(\|x - \xi\|^2) \langle x - \xi, C^0 \xi \rangle \\ &\quad - \left( \sup_{r \geq 0} \varphi'_2(r) + \cdots + \sup_{r \geq 0} \varphi'_{k+1}(r) \right) \\ (3.26) \quad &\quad + \inf_{\|q\| \leq \varphi'_2(d_C(x, y_1)) + \cdots + \varphi'_{k+1}(d_C(x, y_k))} (\langle F(x), Dg(x) + q \rangle \\ &\quad + \tfrac{1}{2} \|B^*(x)(Dg(x) + q)\|_{U_0}^2). \end{aligned}$$

Then  $\tilde{H}_1 \subset \bar{H}_1$ .

LEMMA 3.5.  $\bar{f}$  is a viscosity subsolution of (3.17) for  $H_0$  if and only if it is a viscosity subsolution of (3.19) for  $\bar{H}_0$ ; both imply  $\bar{f}$  is also a viscosity subsolution of (3.21) for  $\tilde{H}_0$ .

$\underline{f}$  is a viscosity supersolution of (3.18) for  $H_1$  if and only if it is a viscosity supersolution of (3.20) for  $\bar{H}_1$ ; both imply  $\underline{f}$  is also a viscosity supersolution of (3.22) for  $\tilde{H}_1$ .

Hence, the comparison principle for subsolution of (3.21) and supersolution of (3.22) implies the comparison principles for (3.17) and (3.18), as well as those for (3.19) and (3.20).

In this lemma viscosity solution is always meant in the sense of Definition 1.14.

PROOF. The conclusion follows from the fact that  $\tilde{H}_0 \subset \bar{H}_0$ ,  $\tilde{H}_1 \subset \bar{H}_1$ ; and the definition of viscosity solution in Definition 1.14.  $\square$

We discuss some properties enjoyed by functions in  $\mathcal{D}(\tilde{H}_i)$ ,  $i = 0, 1$ .

LEMMA 3.6. Let  $f_0 \in \mathcal{D}(\tilde{H}_0)$ . Suppose  $x_0 \in E$  satisfies  $(\bar{f} - f_0)(x_0) = \sup_{x \in E} (\bar{f} - f_0)(x)$ . Let  $0 \leq \varphi \in C^2([0, \infty))$  be nondecreasing,  $\varphi(r) = r$  when  $r \leq 1$  and  $\varphi(r) = 2$  when  $r \geq 2$ . Let  $\theta > 0$ . We introduce perturbation of  $f_0$ :

$$(3.27) \quad f_\theta(x) = f_0(x) + \theta \varphi(d_C(x, x_0)).$$

Then  $f_\theta$  has the following properties:

(a)  $f_\theta \in \mathcal{D}(\tilde{H}_0)$  and

$$(\bar{f} - f_\theta)(x_0) > (\bar{f} - f_\theta)(x), \quad x \neq x_0.$$



(b) For any  $\{x_n\} \subset E$  satisfying

$$\lim_{n \rightarrow \infty} (\bar{f} - f_\theta)(x_n) = \sup_{x \in E} (\bar{f} - f_\theta)(x),$$

we have  $x_n \rightarrow x_0$  and  $\bar{f}(x_n) \rightarrow \bar{f}(x_0)$ .

(c)

$$\limsup_{\theta \rightarrow 0+} (\tilde{H}_0 f_\theta)^*(x_0) \leq (\tilde{H}_0 f_0)^*(x_0),$$

$$\liminf_{\theta \rightarrow 0+} (\tilde{H}_1 f_\theta)_*(x_0) \geq (\tilde{H}_1 f_0)_*(x_0).$$

PROOF. Part (a) follows from the definition of  $f_\theta$ .

We prove part (b) next. By (a),

$$\lim_n (\bar{f} - f_\theta)(x_n) = \sup_x (\bar{f} - f_\theta)(x) = (\bar{f} - f_\theta)(x_0) = (\bar{f} - f_0)(x_0).$$

Therefore

$$\begin{aligned} (\bar{f} - f_0)(x_0) &= (\bar{f} - f_\theta)(x_0) = \lim_n (\bar{f} - f_\theta)(x_n) \\ &= \lim_n (\bar{f} - f_0)(x_n) - \theta \varphi(d_C(x_n, x_0)) \\ &\leq \liminf_n (\bar{f} - f_0)(x_n) = (\bar{f} - f_0)(x_0). \end{aligned}$$

Hence

$$\begin{aligned} \lim_{n \rightarrow \infty} \theta d_C(x_n, x_0) &= \lim_{n \rightarrow \infty} \{(\bar{f} - f_0)(x_n) - ((\bar{f} - f_0)(x_n) - \theta \varphi(d_C(x_n, x_0)))\} \\ &= 0, \end{aligned}$$

which implies  $x_n \rightarrow x_0$  and  $(\bar{f} - f_0)(x_n) \rightarrow (\bar{f} - f_0)(x_0)$ .

Part (c) follows from direct verification.  $\square$

LEMMA 3.7.

- (a) If  $\bar{f}$  is a viscosity subsolution of (3.21) in the sense of Definition 1.14, then it is also a viscosity subsolution in the sense of Definition 1.16 for  $\tilde{H}_0$ .
- (b) If  $\underline{f}$  is a viscosity supersolution of (3.22) in the sense of Definition 1.14, then it is also a viscosity supersolution in the sense of Definition 1.16 for  $\tilde{H}_1$ .

PROOF. We prove part (a) only. The proof for part (b) is similar. Let  $\varphi_i \in \mathcal{T}$  [see (1.32)],  $y_i \in E$  and  $\xi \in \mathcal{D}(C)$ . Consider

$$g_0(x) = \varphi_1(\|x - \xi\|^2),$$

$$f_0(x) = g_0(x) + \varphi_2(d_C(x, y_1)) + \cdots + \varphi_{k+1}(d_C(x, y_k)) \in \mathcal{D}(\tilde{H}_0),$$

and  $x_0 \in E$  such that  $(\bar{f} - f_0)(x_0) = \sup_{x \in E} (\bar{f} - f_0)(x)$ . Define  $f_\theta$  according to (3.27). By Lemma 3.6,  $f_\theta \in \mathcal{D}(\tilde{H}_0)$  and

$$(\bar{f} - f_\theta)(x_0) > (\bar{f} - f_\theta)(x), \quad x \neq x_0.$$

Since  $\bar{f}$  is a subsolution of (3.21) in the sense of Definition 1.14, there exists a sequence  $\{x_n\} \subset E$  such that

$$\lim_n (\bar{f} - f_\theta)(x_n) = \sup_x (\bar{f} - f_\theta)(x)$$

and

$$\limsup_{n \rightarrow \infty} (\alpha^{-1}(\bar{f} - h)(x_n) - (\tilde{H}_0 f_\theta)^*(x_n)) \leq 0.$$

By Lemma 3.6,  $x_n \rightarrow x_0$ ,  $\bar{f}(x_n) \rightarrow \bar{f}(x_0)$  and

$$\limsup_{\theta \rightarrow 0+} (\tilde{H}_0 f_\theta)^*(x_0) \leq (\tilde{H}_0 f_0)^*(x_0).$$

Hence

$$\alpha^{-1}(\bar{f} - h)(x_0) \leq \limsup_{\theta \rightarrow 0+} (\tilde{H}_0 f_\theta)^*(x_0) \leq (\tilde{H}_0 f_0)^*(x_0). \quad \square$$

We define  $\hat{H}_0$  and  $\hat{H}_1$  below. For

$$(3.28) \quad \begin{aligned} g(x) &= \frac{\mu}{2} \|x - \xi\|^2 & \forall \mu > 0, \xi \in \mathcal{D}(C), \\ f(x) &= g(x) + \rho d_C(x, y) & \forall y \in E, \rho > 0 \end{aligned}$$

(recall the definition of  $d_C$  in Definition 1.13), we define

$$(3.29) \quad \begin{aligned} \hat{H}_0 f(x) &= \mu \langle x - \xi, C^0 \xi \rangle + \rho \\ &+ \sup_{\|q\| \leq \rho} (\langle F(x), Dg(x) + q \rangle + \frac{1}{2} \|B^*(x)(Dg(x) + q)\|_{U_0}^2). \end{aligned}$$

Similarly, for

$$(3.30) \quad \begin{aligned} g(x) &= -\frac{\mu}{2} \|x - \xi\|^2 & \forall \mu > 0, \xi \in \mathcal{D}(C), \\ f(x) &= g(x) - \rho d_C(x, y) & \forall y \in E, \rho > 0, \end{aligned}$$

we define

$$(3.31) \quad \begin{aligned} \hat{H}_1 f(x) &= -\mu \langle x - \xi, C^0 \xi \rangle - \rho \\ &+ \inf_{\|q\| \leq \rho} (\langle F(x), Dg(x) + q \rangle + \frac{1}{2} \|B^*(x)(Dg(x) + q)\|_{U_0}^2). \end{aligned}$$

$\tilde{H}_0, \tilde{H}_1$  are local operators, therefore we can get rid of the localization functions  $\varphi_k \in \mathcal{T}$  to arrive at  $\hat{H}_0, \hat{H}_1$ . We omit the proof here. Such argument is standard. In Lemma 6.1 and Theorem 6.1 of [20], these types of arguments are used to prove equivalence of different definitions of viscosity solution. We have the following conclusion.

LEMMA 3.8. *If  $\bar{f}$  is the viscosity subsolution of (3.21) for  $\tilde{H}_0$ , then it is also the subsolution of (3.23) for  $\hat{H}_0$ ; both in the sense of Definition 1.16.*

*If  $\underline{f}$  is the viscosity supersolution of (3.22) for  $\tilde{H}_1$ , then it is also the supersolution of (3.24) for  $\hat{H}_1$ ; both in the sense of Definition 1.16.*

Summarizing conclusions in Lemmas 3.5, 3.7 and 3.8, we have the next result.

LEMMA 3.9. *Let  $\bar{f}$  be a subsolution to (3.17) for  $H_0$  in the sense of Definition 1.14 (Feng and Kurtz); then it is a subsolution to (3.23) for  $\hat{H}_0$  in the sense of Definition 1.16 (Tataru–Crandall–Lions).*

*Let  $\underline{f}$  be a supersolution to (3.18) for  $H_1$  in the sense of Definition 1.14 (Feng and Kurtz); then it is a supersolution to (3.24) for  $\hat{H}_1$  in the sense of Definition 1.16 (Tataru–Crandall–Lions).*

We will study the comparison principle for viscosity solutions in the sense of Definition 1.16 in Section 5. In view of Lemma 3.9, Theorem 5.1 implies the following.

LEMMA 3.10. *Let  $\bar{f}$  be a subsolution to (3.17) for  $H_0$  and let  $\underline{f}$  be a supersolution to (3.18) for  $H_1$ , both in the sense of Definition 1.14. Then  $\bar{f} \leq \underline{f}$ .*

### 3.5. The large deviation theorem.

THEOREM 3.11. *Suppose Conditions 1.11 and 1.12 are satisfied. Let  $X_n \in C_E[0, \infty)$  be the solution of (1.7). Suppose further that  $\{X_n(0)\}$  satisfies the large deviation principle with good rate function  $I_0$  on  $E$ .*

*Then:*

- (a)  $\{X_n\}$  is exponentially tight;
- (b) the following limit exists and defines an operator semigroup on  $C_b(E)$ :

$$(3.32) \quad V(t)f(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \log E[e^{nf(X_n(t))} | X_n(0) = x];$$

- (c) the large deviation principle holds for  $\{X_n\}$  with good rate function  $I$ :

$$(3.33) \quad I(x) = I_0(x(0)) + \sup_{0 \leq t_1 \leq \dots \leq t_m} \left( \sum_{i=1}^m I_{t_i - t_{i-1}}(x(t_i) | x(t_{i-1})) \right),$$

where

$$I_t(y|x) = \sup_{f \in C_b(E)} (f(y) - V(t)f(x)).$$

PROOF. Define  $F \subset C_b(E)$  according to (3.12). The operator convergence in Lemma 3.4 and the comparison principle in Lemma 3.10 imply that Theorem 2.3 holds. Consequently the conclusion follows.  $\square$

**4. Application to special cases.** We solve Examples 1.2, 1.5 and 1.8 as special cases of Theorem 3.11.

4.1. *Stochastic Allen-Cahn equation.* Recall that we take  $E = U_0 = L^2(\mathcal{O})$ . Let  $\omega = \sup_r |V''(r)|$ . We take

$$(4.1) \quad (Cx)(\theta) \equiv \Delta x(\theta) - V'(x(\theta)) + V'(0) - \omega x(\theta)$$

where

$$\mathcal{D}(C) = H^2(\mathcal{O}) \equiv \left\{ x : x, \frac{\partial}{\partial \theta_i} x, \frac{\partial^2}{\partial \theta_i \partial \theta_j} x \in L^2(\mathcal{O}), i, j = 1, \dots, d \right\}$$

and

$$F(x) = -V'(0) + \omega x.$$

$C$  is  $m$ -dissipative by the usual theory of semilinear equation. Recall the  $C_n$  and  $F_n$  in (1.18); we have

$$(4.2) \quad \begin{aligned} \lim_{n \rightarrow \infty} \|C_n \xi - C \xi\|_{L^2(\mathcal{O})} &= 0 \quad \forall \xi \in \mathcal{D}(C), \\ \lim_{x_n \rightarrow x_0} F_n(x_n) &= F(x_0). \end{aligned}$$

In addition, let  $B_n(x)$  be defined according to (1.16); then

$$\lim_{x_n \rightarrow x_0, p_n \rightarrow p_0} \|B_n^*(x_n) p_n\|_{L^2(\mathcal{O})} = \|B^*(x_0) p_0\|_{L^2(\mathcal{O})}.$$

Let  $\{e_1, \dots, e_k, \dots\}$  be the orthonormal system for  $U_0 = L^2(\mathcal{O})$  as defined in (1.11). Let  $\sigma, \varphi$  be defined according to (1.9). Since

$$\begin{aligned} \|B_n(x_n)\|_{L^2(U_0, E)}^2 &= \text{Tr}(B_n^*(x_n) B_n(x_n)) = \sum_k \|B_n(x_n) e_k\|^2 \\ &= \sum_k \sum_{i=(1, \dots, 1)}^{(m_n, \dots, m_n)} \langle \varphi(\cdot, \langle P_n x_n, \xi \rangle) e_k, e_i \rangle^2 \\ &= \sum_{i=(1, \dots, 1)}^{(m_n, \dots, m_n)} \sum_k \langle \varphi(\cdot, \langle P_n x_n, \xi \rangle) e_i, e_k \rangle^2 \\ &= \sum_{i=(1, \dots, 1)}^{(m_n, \dots, m_n)} \|\varphi(\cdot, \langle P_n x_n, \xi \rangle) e_i\|^2 \leq m_n^d \sup_{\theta, r} \varphi^2(\theta, r), \end{aligned}$$

Condition 1.11(3) holds under the scaling requirement (1.14).

Finally, Condition 1.12 is verified by Lemma A.5. Therefore, Theorem 1.4 follows from Theorem 3.11.

To simplify the form of the rate function from (1.31) to a time integral form as in (1.19), we need additional work. The basic idea is the same as that presented in Example 2.5, with some technical complications because of the infinite-dimensional state space. A general result for rate function representation is developed in Chapter 8 of [18]. Applying such result, representation (1.19) for lattice versions of the stochastic Allen–Cahn equation is proved rigorously in Chapter 13 of [18]. This procedure can be carried out similarly here. Below, we only provide a sketch.

First, the form of  $\hat{H}_0$ ,  $\hat{H}_1$  in (3.29) and (3.31) induces an optimal controlled PDE problem:

$$(4.3) \quad \begin{aligned} \frac{\partial}{\partial t} x(t, \theta) &= Cx + F(x) + B(x)u(t) \\ &= \Delta x(t, \theta) - V'(x(t, \theta)) + \sigma(x, \theta)u(t, \theta) \end{aligned}$$

which is well defined under a finite *running cost* assumption:

$$\frac{1}{2} \int_0^\infty \int_{\mathcal{O}} u^2(t, \theta) d\theta dt < \infty.$$

By the dynamic programming principle and the comparison principle for (3.23) and (3.24), we can prove that the  $V(t)$  in (3.32) has the form

$$V(t)f(x_0) = \sup \left\{ f(x(t)) - \frac{1}{2} \int_0^t \int_{\mathcal{O}} u^2(s, \theta) d\theta ds : (x, u) \text{ satisfies (4.3) and } x(0) = x_0 \right\}.$$

Then from this, we derive

$$\begin{aligned} I_t(x_1|x_0) &= \inf \left\{ \int_0^t \int_{\mathcal{O}} \frac{1}{2} u^2(t, \theta) d\theta ds \mid \right. \\ &\quad \left. (x, u) \text{ satisfies (4.3) with } x(0) = x_0, x(t) = x_1 \right\}. \end{aligned}$$

Combine the above form with (4.3); the variational form of  $I(x)$  in (1.19) follows.

**4.2. Stochastic Cahn–Hilliard equation.** Let  $\omega \equiv \sup_r |V''(r)|^2/4$ . We define

$$(4.4) \quad (Cx)(\theta) \equiv \Delta(-\Delta x(\theta) + V'(x(\theta))) - \omega x(\theta)$$

for

$$x \in \mathcal{D}(C) = W^{4,2}(\mathcal{O}) \\ \equiv \left\{ x : x, \frac{\partial}{\partial \theta_i} x, \frac{\partial^2}{\partial \theta_i \partial \theta_j} x, \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} x, \frac{\partial^4}{\partial \theta_i \partial \theta_j \partial \theta_k \partial \theta_l} x \in L^2(\mathcal{O}) \right\}$$

and

$$(F(x))(\theta) \equiv \omega x(\theta), \quad x \in L^2(\mathcal{O}).$$

We recall that the  $C_n, F_n$  are defined as in (1.23). By Lemma A.1,  $C$  and  $C_n$  are  $m$ -dissipative in  $L^2(\mathcal{O})$ . Furthermore, the type of convergence in (4.2) holds here by direct verification.

By Lemma A.7, Condition 1.12 is also satisfied.

Theorem 1.7 follows from Theorem 3.11.

The rate function representation in (1.24) can be proved using similar arguments as in the Allen–Cahn case. The controlled PDE becomes

$$\frac{\partial}{\partial t} x(t, \theta) = Cx + F(x) + B(x)u(t) = \Delta(-\Delta x(t, \theta) + V'(x(t, \theta))) + u(t, \theta).$$

The running cost structure is the same.

**4.3. Stochastic quasilinear equation with viscosity.** Let  $C_n$  be defined according to (1.27) and let

$$(4.5) \quad Cx = \alpha \Delta_\theta x - \partial_\theta \phi(x) - \omega x, \quad x \in H^2(\mathcal{O}).$$

Then both  $C_n$  and  $C$  are  $m$ -dissipative operators in  $L^2(\mathcal{O})$  (Lemma A.4).

The compact containment estimate is provided in Lemma A.9. Following the same arguments as above, Theorem 1.9 follows as a special case of Theorem 3.11.

Rate function representation is the same as the Allen–Cahn case. The controlled PDE is

$$\frac{\partial}{\partial t} x(t, \theta) = Cx + F(x) + B(x)u(t) = \alpha \partial_{\theta\theta}^2 x(t, \theta) - \partial_\theta \phi(x(t, \theta)) + u(t, \theta).$$

**5. A class of Hamilton–Jacobi equation in Hilbert space.** The purpose of this section is to present a self-contained proof of the comparison principle for (5.2) and (5.3) in the viscosity solution sense by Tataru–Crandall–Lions (Definition 1.16)—Theorems 5.1 and 5.2. The whole section is independent of the rest of the paper and can be read separately.

We point out that the comparison results in [31, 32] and in [8] cannot be directly borrowed here, because the  $\hat{H}_0, \hat{H}_1$  are not exactly of the same form as considered there. For example, when defining these operators, we restrict the domains and use rougher estimates than the  $D_C^+$  and  $D_C^-$  (Definition 3.3).

This allows us to relate  $\hat{H}_0, \hat{H}_1$  with other operators which arise as the kind of limits required by Theorem 2.3 with graphs contained in  $C_b(E) \times B(E)$ . Second but more importantly, the quadratic nonlinearity in (5.6) is worse than that assumed in (ii) of (49) in [8]. We explore convexity to cure this problem. Despite these differences, the main ideas of [8, 31, 32], still apply and all we need are modifications and refinements at various places. We follow [8] and present the proof through a doubling technique—Lemma 5.9. To make this paper self-contained, we will repeat the important steps of [8] and omit minor details. We make detailed references for all omitted steps so that they can easily be recovered if needed.

Let  $E$  be a separable real Hilbert space, let  $C \subseteq E \times E$  be a possibly multivalued nonlinear  $m$ -dissipative operator with  $\overline{\mathcal{D}(C)} = E$  and  $(0, 0) \in C$ . Further suppose function  $F: E \rightarrow E$  and operator  $B(x): E \rightarrow E$  for each  $x \in E$  satisfy

$$(5.1) \quad L_{F,B} \equiv \sup_{x \neq y} \frac{\|F(x) - F(y)\| + \|B(x) - B(y)\|}{\|x - y\|} < \infty.$$

Recall the definition of  $\hat{H}_0, \hat{H}_1 \subset C(E) \times M(E)$  in (3.29) and (3.31); we consider the following Hamilton–Jacobi equations written in the resolvent form: let  $h \in C_b(E)$  and  $\alpha > 0$ ,

$$(5.2) \quad (I - \alpha \hat{H}_0)f = h$$

and

$$(5.3) \quad (I - \alpha \hat{H}_1)f = h.$$

We prove the following.

**THEOREM 5.1** (Comparison principle). *Let  $\bar{f}$  be a subsolution of (5.2) and  $\underline{f}$  be a supersolution of (5.3), both in the sense of Definition 1.16.*

*Suppose  $h$  is uniformly continuous and (5.1) is satisfied. Then*

$$\bar{f} \leq \underline{f}.$$

Indeed, we will prove a theorem covering more general situations. Let  $G(x, p) \in C(E \times E)$ . We define single-valued operators  $\hat{H}_0, \hat{H}_1 \subset C(E) \times M(E)$ : for each  $f$  in (3.28), we define

$$(5.4) \quad \hat{H}_0 f(x) \equiv \mu \langle x - \xi, C^0 \xi \rangle + \rho + \sup_{\|q\| \leq \rho} G(x, Dg(x) + q).$$

For each  $f$  in (3.30), we define

$$(5.5) \quad \hat{H}_1 f(x) \equiv -\mu \langle x - \xi, C^0 \xi \rangle - \rho + \inf_{\|q\| \leq \rho} G(x, Dg(x) + q).$$

The operators in (5.2) and (5.3) are special cases of the above ones with  $G$  given by

$$(5.6) \quad G(x, p) = \langle F(x), p \rangle + \frac{1}{2} \|B^*(x)p\|_{U_0}^2.$$

**THEOREM 5.2.** *Suppose  $\alpha > 0$ ,  $h$  is uniformly continuous and Condition 5.5 is satisfied for  $G$ . Define  $\hat{H}_0, \hat{H}_1$  according to (5.4) and (5.5).*

*Let  $\bar{f}$  be a subsolution of (5.2) and  $\underline{f}$  be a supersolution of (5.3), both in the sense of Definition 1.16.*

*Then*

$$\bar{f} \leq \underline{f}.$$

**5.1. Perturbed optimization principle.** There is no a priori guarantee that the extrema in (1.40) and (1.41) of Definition 1.16 always exist. We have to carefully choose test functions  $f_0$  to make sure that the definition is not an empty one. Ekeland's perturbed optimization principle [14] claims that, if we add a small perturbation to the test function using the norm of the Hilbert space, we can always attain the extrema. If we apply this technique in the viscosity solution context, we also want the perturbed  $Hf_0$  to be close to the unperturbed one, so that the equation we consider does not change much. These considerations lead to the Tataru distance function  $d_C$  in Definition 1.13.

The following is adapted from Proposition 2.1 of [31], which generalizes Ekeland's principle. In the adaptation, we have taken  $g = -u$ , where  $u$  is the function in the original proposition.

**LEMMA 5.3.** *Let  $K$  be an abstract set and  $B: K \times K \rightarrow [0, +\infty)$ , with the following properties:*

- (a)  $B(x, x) = 0$  for all  $x \in K$ ;
- (b)  $B(x, y) + B(y, z) \geq B(x, z)$  for all  $x, y, z \in K$ ;
- (c) for each  $\{x_n\} \subset K$  satisfying  $\sum_{n=1}^{\infty} B(x_n, x_{n+1}) < \infty$ , there exists  $x \in K$  such that  $\lim_{n \rightarrow +\infty} B(x_n, x) = 0$ .

*Let  $g: K \rightarrow [-\infty, +\infty)$ ,  $\sup_{x \in K} g(x) < +\infty$ . Furthermore, if  $\{x_n\} \subset K$  and  $x \in K$  satisfy  $\sum_{n=1}^{\infty} B(x_n, x_{n+1}) < +\infty$  and  $\lim_{n \rightarrow +\infty} B(x_n, x) = 0$ , then  $g(x) \geq \limsup g(x_n)$ .*

*Then, for each  $\varepsilon > 0$ , and  $x_0$  such that  $g(x_0) \neq -\infty$ , there exists  $x_\varepsilon$  such that:*

- (1)  $g(x_0) + \varepsilon B(x_0, x_\varepsilon) \leq g(x_\varepsilon)$ ,
- (2)  $g(x) - \varepsilon B(x, x_\varepsilon) \leq g(x_\varepsilon)$ ,  $x \in K$ .



REMARK 5.4. Let function  $g: E \rightarrow R$  be bounded and upper semicontinuous in the norm topology of  $E$ . If we take  $K = E$  and  $B = d_C$ , then the assumptions regarding  $B$  are satisfied (Proposition 2.2 of [31]), and:

(a) For each  $x_0 \in E$  and  $\varepsilon > 0$ , there exists an  $x_1 \in E$  such that

$$(5.7) \quad g(x_0) + \varepsilon d_C(x_0, x_1) \leq g(x_1)$$

and

$$(5.8) \quad g(x) - \varepsilon d_C(x, x_1) \leq g(x_1), \quad x \in E.$$

(b) Let  $\varepsilon > 0$  and  $x_0 \in E$  be such that

$$\sup_{x \in E} g(x) \leq g(x_0) + \varepsilon^2.$$

Then there exists  $x_1 \in E$  such that not only (5.7) and (5.8) hold, but also

$$d_C(x_0, x_1) \leq \varepsilon.$$

Part (a) is a consequence of the above proposition. See also Lemmas 2.4, 2.5 of [8]. Part (b) follows from (5.7):

$$\varepsilon d_C(x_0, x_1) \leq g(x_1) - g(x_0) \leq g(x_1) - \sup_{x \in E} g(x) + \varepsilon^2 \leq \varepsilon^2.$$

Similarly, an analogous result holds when we take  $K = E \times E$  and  $B((x_1, y_1), (x_2, y_2)) = d_C(x_1, x_2) + d_C(y_1, y_2)$ . We will need such result for (5.14).

5.2. *The comparison principle.* We make the following structural assumption about  $G$ .

CONDITION 5.5.

(1)  $G(x, p) \in C(E \times E)$ ; for each  $\lambda > 1$  and  $M > 0$  fixed, there exist  $\rho_\lambda(r)$ ,  $\sigma_M(r) \in C(R^+)$  with  $\rho_\lambda(0) = 0$  and  $\sigma_M(0) = 0$  such that

$$\lambda G\left(x, \frac{\mu(x-y)}{\lambda}\right) - G(y, \mu(x-y)) \leq \rho_\lambda(\|x-y\| + \mu\|x-y\|^2), \quad \mu > 0,$$

and

$$\sup_{\|x\| + \|p\| + \|q\| < M, \|q-p\| \leq r} |G(x, p) - G(x, q)| \leq \sigma_M(r).$$

(2) There exists a nondecreasing function  $0 \leq \varphi \in C^1([0, \infty))$  slowly growing to infinity in the sense that  $\lim_{r \rightarrow \infty} \varphi(r) = \infty$  and  $\sup_{r \geq 0} |r\varphi'(r^2)| + |\varphi'(r)| < \infty$ .  $\varphi' > 0$ .

For  $M > 0$ ,  $\lambda > 1$ , there exists  $\gamma_{\lambda, M}(r) \in C(R^+)$  with  $\gamma_{\lambda, M}(0) = 0$  such that

$$\sup_{x \in E, \|p\| \leq M} \lambda G\left(x, \frac{p + \varepsilon \varphi'(\|x\|^2)x}{\lambda}\right) - G(x, p) \leq \gamma_{\lambda, M}(\varepsilon).$$

REMARK 5.6. Condition 5.5 implies that for any  $1 < \lambda_0 < \lambda_1$ ,

$$\begin{aligned} \lambda_1 G\left(x, \frac{\mu(x-y)}{\lambda_1}\right) - \lambda_0 G\left(y, \frac{\mu(x-y)}{\lambda_0}\right) \\ \leq \lambda_0 \rho_{\lambda_1/\lambda_0} \left( \|x-y\| + \frac{\mu}{\lambda_0} \|x-y\|^2 \right) \end{aligned}$$

and

$$\sup_{x \in E, \|p\| \leq M} \left( \lambda_1 G\left(x, \frac{p + \varepsilon \varphi'(\|x\|^2)x}{\lambda_1}\right) - \lambda_0 G\left(x, \frac{p}{\lambda_0}\right) \right) \leq \lambda_0 \gamma_{\lambda_1/\lambda_0, M} \left( \frac{\varepsilon}{\lambda_0} \right).$$

$\varphi(r) = \log(1+r)$  satisfies the growth requirement in the condition. For many examples, such choice is good enough.

LEMMA 5.7. Suppose (5.1) is satisfied; then the  $G$  in (5.6) satisfies Condition 5.5.

PROOF. To verify this, let  $\lambda > 1$ ; noting

$$\left( \frac{1}{\lambda} - 1 \right) x^2 + \frac{2}{\lambda} bx + \frac{1}{\lambda} b^2 \leq \frac{b^2}{\lambda - 1},$$

we have

$$\begin{aligned} \lambda G\left(x, \frac{p}{\lambda}\right) - G(y, p) &= \frac{1}{2} \left( \frac{1}{\lambda} \|B^*(x)p\|_{U_0}^2 - \|B^*(y)p\|_{U_0}^2 \right) + \langle F(x) - F(y), p \rangle \\ &\leq \frac{1}{2} \left( \frac{1}{\lambda} \|(B^*(x) - B^*(y))p\|_{U_0}^2 + \left( \frac{1}{\lambda} - 1 \right) \|B^*(y)p\|_{U_0}^2 \right. \\ &\quad \left. + \frac{2}{\lambda} \|(B^*(x) - B^*(y))p\|_{U_0} \|B^*(y)p\|_{U_0} \right) \\ &\quad + L_{F,B} \|x - y\| \|p\| \\ &\leq \frac{1}{2} \frac{\|(B^*(x) - B^*(y))p\|_{U_0}^2}{\lambda - 1} + L_{F,B} \|x - y\| \|p\| \\ &\leq \frac{1}{2} \frac{(L_{F,B} \|x - y\| \|p\|)^2}{\lambda - 1} + L_{F,B} \|x - y\| \|p\|, \end{aligned}$$

where  $L_{F,B}$  is the one in (5.1). We can take

$$\rho_\lambda(r) = \frac{1}{2} \frac{L_{F,B}^2}{\lambda - 1} r^2 + L_{F,B} r.$$

Similarly, denoting  $C_0 = \|B(0)\|$ ,

$$\lambda G\left(x, \frac{p + \varepsilon q}{\lambda}\right) - G(x, p) \leq \frac{1}{2} \left( \left( \frac{1}{\lambda} - 1 \right) \|B^*(x)p\|_{U_0}^2 + \frac{1}{\lambda} \|B^*(x)\varepsilon q\|_{U_0}^2 \right)$$

$$\begin{aligned}
& + \frac{2}{\lambda} \|B^*(x)p\|_{U_0} \|B^*(x)\varepsilon q\|_{U_0} \Big) \\
& + \varepsilon \|F(x)\| \|q\| \\
& \leq \frac{1}{2} \frac{\|B^*(x)\varepsilon q\|_{U_0}^2}{\lambda - 1} + \varepsilon L_{F,B} \|x\| \|q\| \\
& \leq \frac{\varepsilon^2 (C_0 \|q\| + L \|q\| \|x\|)^2}{2(\lambda - 1)} + \varepsilon L_{F,B} \|x\| \|q\|.
\end{aligned}$$

We can take  $\varphi(r) = \log(1 + r)$  and

$$\gamma_{\lambda,M}(\varepsilon) = \frac{\varepsilon^2 (C_0 + L)^2}{2(\lambda - 1)} + \varepsilon L_{F,B}.$$

Finally, for each  $M > 0$ , we denote

$$\begin{aligned}
C_1(M) &= \sup_{\|x\| < M} \|B(x)\|, \\
C_2(M) &= \sup_{\|x\| < M} \|B(x)\|^2 M, \\
C_3(M) &= \sup_{\|x\| < M} \|F(x)\|.
\end{aligned}$$

Then for  $\|x\| + \|p\| + \|q\| < M$ ,

$$\begin{aligned}
|G(x, p) - G(x, q)| &\leq (\|B^*(x)(p - q)\|_{U_0}^2 \\
&\quad + 2\|B^*(x)q\|_{U_0} \|B^*(x)(p - q)\|_{U_0} + \|F(x)\| \|p - q\|) \\
&\leq C_1^2(M) \|p - q\|^2 + 2C_2(M) \|p - q\| + C_3(M) \|p - q\|.
\end{aligned}$$

We can take  $\sigma_M(r) = C_1^2(M)r^2 + (2C_2(M) + C_3(M))r$ .  $\square$

Now, we prove Theorem 5.2 in several steps.

We endow the product space  $E \times E$  with inner product

$$\langle (x, y), (\xi, \eta) \rangle = \langle x, \xi \rangle + \langle y, \eta \rangle.$$

Denote

$$\mathcal{C}(x, y) = (Cx, Cy).$$

By the  $m$ -dissipativity of  $C$ ,  $\mathcal{C}$  induces a semigroup:

$$\mathcal{S}(t)(x, y) = (S(t)x, S(t)y), \quad (x, y) \in E \times E.$$

For  $\lambda > 0$ ,  $(x, y), (p, q) \in E \times E$ , we define

$$G_\lambda((x, y); (p, q)) \equiv \lambda G\left(x, \frac{p}{\lambda}\right) - G(y, -q),$$

and define a single-valued operator  $H_{2,\lambda} \subset C(E \times E) \times M(E \times E)$  next. Let  $\mathcal{D} = \mathcal{D}(H_{2,\lambda})$  consist of functions  $\Phi$  defined as follows:

$$(5.9) \quad \phi(x, y) = \frac{\mu}{2} \|x - y\|^2 + \frac{\gamma}{2} (\varphi(\|x\|^2) + \varphi(\|y\|^2)), \quad \mu, \gamma > 0,$$

where  $\varphi \in C^1([0, \infty))$ ,  $\varphi' \geq 0$ ,  $\lim_{r \rightarrow +\infty} \varphi(r) = +\infty$ ; and

$$\Phi(x, y) = \phi(x, y) + \rho(d_C(x, x_0) + d_C(y, y_0)), \quad \rho > 0, x_0, y_0 \in E.$$

We define, for each  $\Phi$ ,

$$(5.10) \quad \begin{aligned} H_{2,\lambda}\Phi(x, y) &= D_C^+\Phi(x, y) \\ &+ \sup_{\|p\|^2 + \|q\|^2 \leq 2\rho^2} G_\lambda((x, y); D\phi(x, y) + (p, q)). \end{aligned}$$

We note that  $H_{2,\lambda}\Phi(x, y)$  may be an unbounded function on  $E \times E$ .

We have a perturbation result.

LEMMA 5.8. *Let  $u \in B(E \times E)$  be upper semicontinuous,  $\Phi_0(x, y) \in \mathcal{D}$  and  $(\hat{x}, \hat{y}) \in E \times E$  satisfy  $(u - \Phi_0)(\hat{x}, \hat{y}) = \sup_{(x, y) \in E \times E} (u - \Phi_0)(x, y)$ . Suppose  $\kappa > 0$ ; we define*

$$(5.11) \quad \Phi_\kappa(x, y) = \Phi_0(x, y) + \kappa(d_C(x, \hat{x}) + d_C(y, \hat{y})).$$

Then  $\Phi_\kappa$  has the following properties:

(a)

$$(u - \Phi_\kappa)(\hat{x}, \hat{y}) > (u - \Phi_\kappa)(x, y), \quad (x, y) \neq (\hat{x}, \hat{y}).$$

(b) For any  $\{(x_n, y_n)\} \subset E \times E$  satisfying

$$\lim_{n \rightarrow \infty} (u - \Phi_\kappa)(x_n, y_n) = \sup_{(x, y) \in E \times E} (u - \Phi_\kappa)(x, y),$$

we have  $(x_n, y_n) \rightarrow (\hat{x}, \hat{y})$  and  $u(x_n, y_n) \rightarrow u(\hat{x}, \hat{y})$ .

PROOF. The same arguments as in the proof of Lemma 3.6 apply here.

□

LEMMA 5.9 (Doubling lemma). *Let  $\lambda > 0$  and Condition 5.5 be satisfied. Define  $u(x, y) = \lambda \bar{f}(x) - \underline{f}(y)$  and  $v(x, y) = \lambda h(x) - h(y)$ , where the  $\bar{f}$  and  $\underline{f}$  are the ones given in Theorem 5.2.*

*Then  $u$  is a viscosity subsolution (in the sense of Definition 1.16) of*

$$(I - \alpha H_{2,\lambda})u = v.$$

PROOF. Let  $\Phi_0 \in \mathcal{D}$  and  $(\hat{x}, \hat{y}) \in E \times E$  satisfy

$$(u - \Phi_0)(\hat{x}, \hat{y}) = \sup_{x, y \in E} (u - \Phi_0)(x, y).$$

We want to show that

$$(5.12) \quad \alpha^{-1}(u(\hat{x}, \hat{y}) - v(\hat{x}, \hat{y})) \leq (H_{2,\lambda}\Phi_0)^*(\hat{x}, \hat{y}).$$

We may assume  $\Phi_0$  takes the following form:

$$\Phi_0(x, y) = \phi(x, y) + \rho(d_C(x, x_0) + d_C(y, y_0))$$

for some  $x_0, y_0 \in E$ , where

$$\phi(x, y) = \frac{\mu}{2}(\|x - y\|^2) + \frac{\gamma}{2}(\varphi(\|x\|^2) + \varphi(\|y\|^2))$$

takes the form in (5.9). Fix  $\kappa > 0$ ; we let

$$\Phi(x, y) = \Phi_0(x, y) + \kappa(d_C(x, \hat{x}) + d_C(y, \hat{y})).$$

By Lemma 5.8,

$$(u - \Phi)(\hat{x}, \hat{y}) > (u - \Phi)(x, y) \quad \forall (x, y) \neq (\hat{x}, \hat{y}).$$

We define

$$\Psi(x, y) = u(x, y) - \Phi(x, y),$$

$$\Psi_\varepsilon(x, y, \xi, \eta) = u(x, y) - \Phi(\xi, \eta) - \frac{1}{2\varepsilon}(\|x - \xi\|^2 + \|y - \eta\|^2)$$

and

$$\begin{aligned} \Psi_{\varepsilon,\delta}(x, y, \xi, \eta) &= u(x, y) - \Phi(\xi, \eta) \\ &\quad - \frac{1}{2\varepsilon}(\|x - \xi\|^2 + \|y - \eta\|^2) - \delta(\|C^0\xi\| + \|C^0\eta\|), \end{aligned}$$

where  $\|C^0\xi\| = +\infty$ , if  $\xi \notin \mathcal{D}(C)$ .

We write the maximum of each of these functions:

$$M = \sup_{x, y \in E} \Psi(x, y) = \Psi(\hat{x}, \hat{y}),$$

$$M_\varepsilon = \sup_{x, y, \xi, \eta \in E} \Psi_\varepsilon(x, y, \xi, \eta),$$

$$M_{\varepsilon,\delta} = \sup_{x, y, \xi, \eta \in E} \Psi_{\varepsilon,\delta}(x, y, \xi, \eta).$$

It follows that  $M \leq M_\varepsilon$ ,  $M_\varepsilon \geq M_{\varepsilon,\delta}$  and  $M_\varepsilon \downarrow M$ , as  $\varepsilon \downarrow 0$  and  $M_{\varepsilon,\delta} \uparrow M_\varepsilon$  as  $\delta \downarrow 0$ . See, for example, page 83 of [8]. The definitions of  $M_\varepsilon, M_{\varepsilon,\delta}$  in [8] are slightly different than here, in the sense that suprema are taken locally for a ball of size  $2r$  with arbitrary  $r > 0$ , instead of over the whole space.

Note that  $\sup_{x,y} |u(x,y)| < +\infty$ ; note also that the form of  $\phi$  in (5.9) implies that  $\lim_{\|x\|+\|y\|\rightarrow+\infty} \Phi(x,y) = +\infty$ , hence the suprema over the whole space are equal to the corresponding suprema over a sufficiently large open ball. Therefore, the same proof in [8] still works here.

- (1) For each  $\varepsilon, \theta > 0$ , we can choose  $\delta = \delta(\varepsilon, \theta)$  and  $(x_{\varepsilon,\theta}, y_{\varepsilon,\theta}, \xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta})$  such that  $\delta(\varepsilon, \theta) \downarrow 0$  as  $\varepsilon, \theta \downarrow 0$  and

$$(5.13) \quad M_\varepsilon - \theta \leq \Psi_{\varepsilon,\delta(\varepsilon,\theta)}(x_{\varepsilon,\theta}, y_{\varepsilon,\theta}, \xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta}).$$

By Lemma 5.3, we can always select the  $(x_{\varepsilon,\theta}, y_{\varepsilon,\theta}, \xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta})$  so that (note that  $\|C^0 x\|$  as a function in  $x$  is lower semicontinuous on  $E$ ; see, e.g., Lemma 2.18 in [26])

$$(5.14) \quad \begin{aligned} & \Psi_{\varepsilon,\delta(\varepsilon,\theta)}(x, y, \xi, \eta) \\ & - \varepsilon(d_C(x, x_{\varepsilon,\theta}) + d_C(y, y_{\varepsilon,\theta}) + d_C(\xi, \xi_{\varepsilon,\theta}) + d_C(\eta, \eta_{\varepsilon,\theta})) \\ & \leq \Psi_{\varepsilon,\delta(\varepsilon,\theta)}(x_{\varepsilon,\theta}, y_{\varepsilon,\theta}, \xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta}) \quad \forall x, y, \xi, \eta \in E. \end{aligned}$$

Let  $x = x_{\varepsilon,\theta}, y = y_{\varepsilon,\theta}$ ; then

$$(5.15) \quad \begin{aligned} & \Phi(\xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta}) + \frac{1}{2\varepsilon}(\|x_{\varepsilon,\theta} - \xi_{\varepsilon,\theta}\|^2 + \|y_{\varepsilon,\theta} - \eta_{\varepsilon,\theta}\|^2) \\ & + \delta(\varepsilon, \theta)(\|C^0 \xi_{\varepsilon,\theta}\| + \|C^0 \eta_{\varepsilon,\theta}\|) \\ & \leq \Phi(\xi, \eta) + \frac{1}{2\varepsilon}(\|x_{\varepsilon,\theta} - \xi\|^2 + \|y_{\varepsilon,\theta} - \eta\|^2) \\ & + \delta(\varepsilon, \theta)(\|C^0 \xi\| + \|C^0 \eta\|) + \varepsilon(d_C(\xi, \xi_{\varepsilon,\theta}) + d_C(\eta, \eta_{\varepsilon,\theta})). \end{aligned}$$

From  $\Psi_\varepsilon \geq \Psi_{\varepsilon,\delta(\varepsilon,\theta)}$  and (5.13),

$$\Psi_\varepsilon(x, y, \xi, \eta) \leq \Psi_\varepsilon(x_{\varepsilon,\theta}, y_{\varepsilon,\theta}, \xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta}) + \theta.$$

Take  $x = x_{\varepsilon,\theta}, y = y_{\varepsilon,\theta}$ , therefore

$$(5.16) \quad \begin{aligned} & \theta + \Phi(\xi, \eta) + \frac{1}{2\varepsilon}(\|x_{\varepsilon,\theta} - \xi\|^2 + \|y_{\varepsilon,\theta} - \eta\|^2) \\ & \geq \Phi(\xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta}) \\ & + \frac{1}{2\varepsilon}(\|x_{\varepsilon,\theta} - \xi_{\varepsilon,\theta}\|^2 + \|y_{\varepsilon,\theta} - \eta_{\varepsilon,\theta}\|^2), \quad \xi, \eta \in E. \end{aligned}$$

- (2) It follows from (5.13) that  $\lim_{\varepsilon \rightarrow 0+, \theta \rightarrow 0+} \Psi_{\varepsilon,\delta(\varepsilon,\theta)}(x_{\varepsilon,\theta}, y_{\varepsilon,\theta}) = M$ , hence

$$(5.17) \quad \begin{aligned} & (x_{\varepsilon,\theta}, y_{\varepsilon,\theta}), (\xi_{\varepsilon,\theta}, \eta_{\varepsilon,\theta}) \rightarrow (\hat{x}, \hat{y}), \\ & \bar{f}(x_{\varepsilon,\theta}) \rightarrow \bar{f}(\hat{x}), \quad \underline{f}(x_{\varepsilon,\theta}) \rightarrow \underline{f}(\hat{y}) \quad \text{as } \varepsilon, \theta \downarrow 0 \end{aligned}$$

(e.g., Step 3 on page 84 of [8] and Lemma 5.8).

Take  $\xi = \xi_{\varepsilon, \theta}$ ,  $\eta = \eta_{\varepsilon, \theta}$ ,  $y = y_{\varepsilon, \theta}$  in (5.14):

$$x \rightarrow \lambda \bar{f}(x) - \frac{1}{2\varepsilon} \|x - \xi_{\varepsilon, \theta}\|^2 - \varepsilon d_C(x, x_{\varepsilon, \theta})$$

has a maximum at  $x_{\varepsilon, \theta}$ . Since  $\bar{f}$  is a subsolution of (5.2) in the sense of Definition 1.16,

$$\begin{aligned} \lambda \frac{\bar{f} - h}{\alpha}(x_{\varepsilon, \theta}) &\leq \frac{1}{\varepsilon} \langle x_{\varepsilon, \theta} - \xi_{\varepsilon, \theta}, C^0 \xi_{\varepsilon, \theta} \rangle + \varepsilon \\ &\quad + \sup_{\|p\| \leq \varepsilon} \lambda G\left(x_{\varepsilon, \theta}, \frac{1}{\lambda} \left( \frac{x_{\varepsilon, \theta} - \xi_{\varepsilon, \theta}}{\varepsilon} + p \right)\right). \end{aligned}$$

Similarly, noting  $y_{\varepsilon, \theta}$  is a maximum of

$$y \rightarrow -\underline{f}(y) - \frac{1}{2\varepsilon} \|y - \eta_{\varepsilon, \theta}\|^2 - \varepsilon d_C(y, y_{\varepsilon, \theta}),$$

by the supersolution property of  $\underline{f}$ ,

$$\begin{aligned} \frac{\underline{f} - h}{\alpha}(y_{\varepsilon, \theta}) &\geq -\frac{1}{\varepsilon} \langle y_{\varepsilon, \theta} - \eta_{\varepsilon, \theta}, C^0 \eta_{\varepsilon, \theta} \rangle - \varepsilon \\ &\quad + \inf_{\|q\| \leq \varepsilon} G\left(y_{\varepsilon, \theta}, -\frac{y_{\varepsilon, \theta} - \eta_{\varepsilon, \theta}}{\varepsilon} + q\right). \end{aligned}$$

Therefore

$$\begin{aligned} &\alpha^{-1}(u - v)(x_{\varepsilon, \theta}, y_{\varepsilon, \theta}) \\ &\leq \left\langle \frac{1}{\varepsilon} (x_{\varepsilon, \theta} - \xi_{\varepsilon, \theta}, y_{\varepsilon, \theta} - \eta_{\varepsilon, \theta}), (C^0 \xi_{\varepsilon, \theta}, C^0 \eta_{\varepsilon, \theta}) \right\rangle + 2\varepsilon \\ (5.18) \quad &+ \sup_{\|p\|^2 + \|q\|^2 \leq 2\varepsilon^2} \left( \lambda G\left(x_{\varepsilon, \theta}, \frac{1}{\lambda} \left( \frac{x_{\varepsilon, \theta} - \xi_{\varepsilon, \theta}}{\varepsilon} + p \right)\right) \right. \\ &\quad \left. - G\left(y_{\varepsilon, \theta}, -\frac{y_{\varepsilon, \theta} - \eta_{\varepsilon, \theta}}{\varepsilon} + q\right) \right). \end{aligned}$$

(3) Apply Lemma A.8 of [8] to (5.16), and take  $\theta = \varepsilon^3$ . We denote

$$x_\varepsilon = x_{\varepsilon, \varepsilon^3}, \quad y_\varepsilon = y_{\varepsilon, \varepsilon^3}, \quad \xi_\varepsilon = \xi_{\varepsilon, \varepsilon^3}, \quad \eta_\varepsilon = \eta_{\varepsilon, \varepsilon^3}.$$

(a) After some algebra (for details, see page 86 of [8]),

$$(5.19) \quad \limsup_{\varepsilon \rightarrow 0+} \left\| D\phi(\hat{x}, \hat{y}) - \left( \frac{x_\varepsilon - \xi_\varepsilon}{\varepsilon}, \frac{y_\varepsilon - \eta_\varepsilon}{\varepsilon} \right) \right\| \leq \rho + \kappa.$$

- (b) Take  $\theta = \varepsilon^3, \xi = S(h)\xi_\varepsilon, \eta = S(h)\eta_\varepsilon$  in (5.15); noting  $\|C^0 S(h)\xi_\varepsilon\| \leq \|C^0 \xi_\varepsilon\|$  and  $\|C^0 S(h)\eta_\varepsilon\| \leq \|C^0 \eta_\varepsilon\|$ ,

$$\begin{aligned} & \frac{1}{2\varepsilon} \frac{1}{h} (\|(x_\varepsilon, y_\varepsilon) - (\xi_\varepsilon, \eta_\varepsilon)\|^2 - \|(x_\varepsilon, y_\varepsilon) - \mathcal{S}(h)(\xi_\varepsilon, \eta_\varepsilon)\|^2) \\ & \leq \frac{1}{h} (\Phi(S(h)\xi_\varepsilon, S(h)\eta_\varepsilon) - \Phi(\xi_\varepsilon, \eta_\varepsilon)) \\ & \quad + \varepsilon \frac{1}{h} (d_C(S(h)\xi_\varepsilon, \xi_\varepsilon) + d_C(S(h)\eta_\varepsilon, \eta_\varepsilon)). \end{aligned}$$

Send  $h \rightarrow 0+$ ; by (3.2) and (3.4),

$$\begin{aligned} & \left\langle \frac{(x_\varepsilon, y_\varepsilon) - (\xi_\varepsilon, \eta_\varepsilon)}{\varepsilon}, (C^0 \xi_\varepsilon, C^0 \eta_\varepsilon) \right\rangle \\ & \leq \limsup_{h \rightarrow 0+} \frac{\Phi(S(h)\xi_\varepsilon, S(h)\eta_\varepsilon) - \Phi(\xi_\varepsilon, \eta_\varepsilon)}{h} + 2\varepsilon. \end{aligned}$$

Hence by (5.17),

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0+} \left\langle \frac{(x_\varepsilon, y_\varepsilon) - (\xi_\varepsilon, \eta_\varepsilon)}{\varepsilon}, (C^0 \xi_\varepsilon, C^0 \eta_\varepsilon) \right\rangle \\ & \leq \limsup_{h \rightarrow 0+, \varepsilon \rightarrow 0+} \frac{\Phi(S(h)\xi_\varepsilon, S(h)\eta_\varepsilon) - \Phi(\xi_\varepsilon, \eta_\varepsilon)}{h} \\ (5.20) \quad & \leq \limsup_{h \rightarrow 0+, (x, y) \rightarrow (\hat{x}, \hat{y})} \frac{\Phi(S(h)x, S(h)y) - \Phi(x, y)}{h} \\ & \leq D_C^+ \Phi(\hat{x}, \hat{y}) \leq D_C^+ \Phi_0(\hat{x}, \hat{y}) + 2\kappa. \end{aligned}$$

Finally, apply (5.19) and (5.20) to (5.18); noting (5.17),

$$\begin{aligned} (5.21) \quad & \frac{u-v}{\alpha}(\hat{x}, \hat{y}) \leq D_C^+ \Phi_0(\hat{x}, \hat{y}) + 2\kappa \\ & + \sup_{\|p\|^2 + \|q\|^2 \leq 2(\rho + \kappa)^2} G_\lambda((\hat{x}, \hat{y}); D\phi(\hat{x}, \hat{y}) + (p, q)). \end{aligned}$$

Send  $\kappa \rightarrow 0$ ;

$$\frac{u-v}{\alpha}(\hat{x}, \hat{y}) \leq (H_{2, \lambda} \Phi_0)^*(\hat{x}, \hat{y}).$$

□

We discuss a property of  $D_C^+$ .

LEMMA 5.10. *Let  $\Phi(x, y) \in \mathcal{D}$  be such that*

$$\Phi(x, y) = \phi(x, y) + \theta(d_C(x, x_0) + d_C(y, y_0)), \quad x_0, y_0 \in E, \theta > 0,$$



where

$$\phi(x, y) = \frac{\mu}{2}\|x - y\|^2 + \frac{\varepsilon}{2}(\varphi(\|x\|^2) + \varphi(\|y\|^2)),$$

$$\mu > 0, \varphi \in C^1([0, \infty)), \varphi' \geq 0.$$

Then

$$(5.22) \quad D_C^+ \phi(x, y) \leq 0$$

and

$$(5.23) \quad D_C^+ \Phi(x, y) \leq 2\theta.$$

PROOF. Equation (5.22) follows because  $S(t)$  is a contraction semigroup with the property  $\|S(t)x\| \leq \|x\|$ :

$$\phi(S(h)x, S(h)y) \leq \phi(x, y) \quad \forall x, y \in E, h > 0.$$

For each  $x, x_0 \in E$ , by the definition of  $d_C$ , there exists  $t_0 > 0$  such that

$$\begin{aligned} d_C(x, x_0) + h &= \inf_{t \geq 0} (t + \|x - S(t)x_0\|) + h = t_0 + \|x - S(t_0)x_0\| + h \\ &\geq t_0 + h + \|S(h)x - S(t_0 + h)x_0\| \geq d_C(S(h)x, x_0) \quad \forall h > 0. \end{aligned}$$

Hence (5.23) follows.  $\square$

PROOF OF THEOREM 5.2. We assume  $(\bar{f} - \underline{f})(z_0) = 4\delta_0 > 0$  for some  $z_0 \in E$  (otherwise, there is nothing to prove). We want to create a contradiction.

Let  $1 < \lambda < 1 + \delta_0/(1 + \sup_x |\bar{f}(x)|)$ . We recall that  $u(x, y) = \lambda\bar{f}(x) - \underline{f}(y)$ . Then  $u(z_0, z_0) \geq 3\delta_0 > 0$ . Let  $\varphi(r)$  be given by Condition 5.5(2); we define

$$\phi(x, y) = \frac{\mu}{2}\|x - y\|^2 + \frac{\varepsilon}{2}(\varphi(\|x\|^2) + \varphi(\|y\|^2)).$$

Hence for  $0 < \varepsilon < (\delta_0)/(1 + \varphi(|z_0|^2))$ ,

$$(5.24) \quad 0 < 2\delta_0 \leq u(z_0, z_0) - \delta_0 \leq (u - \phi)(z_0, z_0) \leq \sup_{x, y \in E} (u - \phi)(x, y).$$

By Lemma 5.3 on perturbed optimization, for each  $\theta > 0$ , there exist  $x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta} \in E$  such that

$$(u - \phi)(x, y) - \theta(d_C(x, x_{\mu, \varepsilon, \theta}) + d_C(y, y_{\mu, \varepsilon, \theta})) \leq (u - \phi)(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta})$$

and

$$(5.25) \quad \sup_{x, y \in E} (u - \phi)(x, y) \leq (u - \phi)(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) + \theta.$$

Denote

$$\Phi(x, y) = \phi(x, y) + \theta(d_C(x, x_{\mu, \varepsilon, \theta}) + d_C(y, y_{\mu, \varepsilon, \theta})),$$

therefore  $u(x, y) - \Phi(x, y)$  attains its maximum at  $(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta})$ . From (5.24) and (5.25),

$$(5.26) \quad \begin{aligned} \delta_0 &\leq \delta_0 + \frac{\mu}{2} \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\|^2 + \frac{\varepsilon}{2} (\varphi(\|x_{\mu, \varepsilon, \theta}\|^2) + \varphi(\|y_{\mu, \varepsilon, \theta}\|^2)) \\ &\leq u(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) \leq \lambda \sup_x |\bar{f}(x)| + \sup_x |\underline{f}(x)| < \infty \end{aligned}$$

for every  $\theta \leq \delta_0$ ,  $|\lambda - 1| \leq \delta_0(1 + \sup_x |\bar{f}(x)|)^{-1}$ . Equation (5.26) implies the existence of constants  $C_\mu, M_\varepsilon$  such that

$$\begin{aligned} \mu \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\| &\leq C_\mu < \infty & \forall \varepsilon > 0, 0 < \theta < \delta_0, \\ \sqrt{\|x_{\mu, \varepsilon, \theta}\|^2 + \|y_{\mu, \varepsilon, \theta}\|^2} &\leq M_\varepsilon < \infty & \forall \mu > 0, 0 < \theta < \delta_0. \end{aligned}$$

In addition, since

$$C_0 \equiv \sup_{r \geq 0} |r\varphi'(r^2)| + |r\varphi'(r)| < \infty,$$

there exists constant  $N_{\varepsilon, \mu}$ ,

$$\begin{aligned} &\|x_{\mu, \varepsilon, \theta}\| + \|y_{\mu, \varepsilon, \theta}\| + \mu \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\| \\ &+ \varepsilon \varphi'(\|x_{\mu, \varepsilon, \theta}\|^2) \|x_{\mu, \varepsilon, \theta}\| + \varepsilon \varphi'(\|y_{\mu, \varepsilon, \theta}\|^2) \|y_{\mu, \varepsilon, \theta}\| \leq N_{\varepsilon, \mu} < \infty, \end{aligned}$$

for  $0 < \theta < \delta_0$ . Since

$$\begin{aligned} D\phi(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) &= (\mu(x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}) + \varepsilon \varphi'(\|x_{\mu, \varepsilon, \theta}\|^2) x_{\mu, \varepsilon, \theta}, \\ &\quad - \mu(x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}) + \varepsilon \varphi'(\|y_{\mu, \varepsilon, \theta}\|^2) y_{\mu, \varepsilon, \theta}), \end{aligned}$$

and  $u$  is a viscosity subsolution of  $(I - \alpha H_{2, \lambda})u = v$  (Lemma 5.9), by the estimate in (5.23) and (5.26), for  $\theta < \delta_0$  and  $|\lambda - 1| \leq \delta_0(1 + 2 \sup_x |h(x)| \vee 2 \sup_x |\bar{f}(x)|)^{-1}$ ,

$$(5.27) \quad \begin{aligned} &\frac{\delta_0}{2} - (h(x_{\mu, \varepsilon, \theta}) - h(y_{\mu, \varepsilon, \theta})) \\ &\leq u(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) - (\lambda h(x_{\mu, \varepsilon, \theta}) - h(y_{\mu, \varepsilon, \theta})) \\ &\leq 2\theta + \sup_{\|p\|^2 + \|q\|^2 \leq 2\theta^2} G_\lambda((x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}); \\ &\quad D\phi(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) + (p, q)) \\ &\leq 2\theta + \sup_{\|p\|^2 + \|q\|^2 \leq 2\theta^2} G_\lambda\{(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}); \\ &\quad (\mu(x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}) \\ &\quad + \varepsilon \varphi'(\|x_{\mu, \varepsilon, \theta}\|^2) x_{\mu, \varepsilon, \theta} + p, \end{aligned}$$

$$\begin{aligned}
& -\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) \\
& + \varepsilon\varphi'(\|y_{\mu,\varepsilon,\theta}\|^2)y_{\mu,\varepsilon,\theta} + q\}.
\end{aligned}$$

We select  $\lambda_0, \lambda_1$  satisfying  $1 < \lambda_0 < \lambda_1 < \lambda$ , and let them be fixed. By Condition 5.5, for  $\|p\| < 1$ ,

$$\begin{aligned}
& \lambda G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) + \varepsilon\varphi'(\|x_{\mu,\varepsilon,\theta}\|^2)x_{\mu,\varepsilon,\theta} + p}{\lambda}\right) \\
& = \lambda G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) + \varepsilon\varphi'(\|x_{\mu,\varepsilon,\theta}\|^2)x_{\mu,\varepsilon,\theta} + p}{\lambda}\right) \\
& \quad - \lambda G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) + \varepsilon\varphi'(\|x_{\mu,\varepsilon,\theta}\|^2)x_{\mu,\varepsilon,\theta}}{\lambda}\right) \\
& \quad + \lambda G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) + \varepsilon\varphi'(\|x_{\mu,\varepsilon,\theta}\|^2)x_{\mu,\varepsilon,\theta}}{\lambda}\right) \\
& \quad - \lambda_1 G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta})}{\lambda_1}\right) + \lambda_1 G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta})}{\lambda_1}\right) \\
& \leq \lambda\sigma_{N_{\varepsilon,\mu}+1}\left(\frac{\|p\|}{\lambda}\right) + \lambda_1\gamma_{\lambda/\lambda_1, C_\mu/\lambda_1}\left(\frac{\varepsilon}{\lambda_1}\right) \\
& \quad + \lambda_1 G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta})}{\lambda_1}\right).
\end{aligned}$$

Similarly, for  $\|q\| < 1$  and  $0 \leq \varepsilon < 1$ ,

$$\begin{aligned}
& G(y_{\mu,\varepsilon,\theta}, \mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) - \varepsilon\varphi'(\|y_{\mu,\varepsilon,\theta}\|^2)y_{\mu,\varepsilon,\theta} - q) \\
& \geq G(y_{\mu,\varepsilon,\theta}, \mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) - \varepsilon\varphi'(\|y_{\mu,\varepsilon,\theta}\|^2)y_{\mu,\varepsilon,\theta}) - \sigma_{N_{\varepsilon,\mu}+1}(\|q\|) \\
& \geq \lambda_0 G\left(y_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta})}{\lambda_0}\right) - \gamma_{\lambda_0, C_\mu+C_0}(\varepsilon) - \sigma_{N_{\varepsilon,\mu}+1}(\|q\|).
\end{aligned}$$

Therefore

$$\begin{aligned}
& G_\lambda\{(x_{\mu,\varepsilon,\theta}, y_{\mu,\varepsilon,\theta}); (\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) + \varepsilon\varphi'(\|x_{\mu,\varepsilon,\theta}\|^2)x_{\mu,\varepsilon,\theta} + p, \\
& \quad - \mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}) + \varepsilon\varphi'(\|y_{\mu,\varepsilon,\theta}\|^2)y_{\mu,\varepsilon,\theta} + q)\} \\
& \leq \lambda_1 G\left(x_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta})}{\lambda_1}\right) - \lambda_0 G\left(y_{\mu,\varepsilon,\theta}, \frac{\mu(x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta})}{\lambda_0}\right) \\
& \quad + \lambda\sigma_{N_{\varepsilon,\mu}+1}\left(\frac{\|p\|}{\lambda}\right) + \lambda_1\gamma_{\lambda/\lambda_1, C_\mu/\lambda_1}\left(\frac{\varepsilon}{\lambda_1}\right) \\
& \quad + \gamma_{\lambda_0, C_\mu+C_0}(\varepsilon) + \sigma_{N_{\varepsilon,\mu}+1}(\|q\|) \\
& \leq \lambda_0\rho_{\lambda_1/\lambda_0}\left(\|x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}\| + \frac{\mu}{\lambda_0}\|x_{\mu,\varepsilon,\theta} - y_{\mu,\varepsilon,\theta}\|^2\right) + \lambda\sigma_{N_{\varepsilon,\mu}+1}\left(\frac{\|p\|}{\lambda}\right)
\end{aligned}$$

$$+ \lambda_1 \gamma_{\lambda/\lambda_1, C_\mu/\lambda_1} \left( \frac{\varepsilon}{\lambda_1} \right) + \gamma_{\lambda_0, C_\mu + C_0}(\varepsilon) + \sigma_{N_{\varepsilon, \mu} + 1}(\|q\|).$$

We rewrite (5.27) next. For  $\theta < \delta_0 < 1$  and  $1 < \lambda \leq 1 + \frac{\delta_0}{1+2\|h\|\sqrt{2}\|\bar{f}\|}$ ,

$$\begin{aligned}
 & \frac{\delta_0}{2} - (h(x_{\mu, \varepsilon, \theta}) - h(y_{\mu, \varepsilon, \theta})) \\
 & \leq 2\theta + \sup_{\|p\|^2 + \|q\|^2 \leq 2\theta^2} \left( \lambda \sigma_{N_{\varepsilon, \mu} + 1} \left( \frac{\|p\|}{\lambda} \right) + \sigma_{N_{\varepsilon, \mu} + 1}(\|q\|) \right) \\
 (5.28) \quad & + \lambda_1 \gamma_{\lambda/\lambda_1, C_\mu/\lambda_1} \left( \frac{\varepsilon}{\lambda_1} \right) + \gamma_{\lambda_0, C_\mu + C_0}(\varepsilon) \\
 & + \lambda_0 \rho_{\lambda_1/\lambda_0} \left( \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\| + \frac{\mu}{\lambda_0} \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\|^2 \right).
 \end{aligned}$$

Let

$$m_\mu = \sup_{x, y \in E} \left( u(x, y) - \frac{\mu}{2} \|x - y\|^2 \right).$$

From (5.25),

$$\begin{aligned}
 u(x, y) - \frac{\mu}{2} \|x - y\|^2 & \leq \liminf_{\varepsilon \rightarrow 0, \theta \rightarrow 0} \left( u(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) - \frac{\mu}{2} \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\|^2 \right) \\
 & \leq \limsup_{\varepsilon \rightarrow 0, \theta \rightarrow 0} \left( u(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) - \frac{\mu}{2} \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\|^2 \right) \\
 & \leq m_\mu, \quad x, y \in E.
 \end{aligned}$$

Hence

$$m_\mu = \lim_{\varepsilon \rightarrow 0, \theta \rightarrow 0} \left( u(x_{\mu, \varepsilon, \theta}, y_{\mu, \varepsilon, \theta}) - \frac{\mu}{2} \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\|^2 \right).$$

Applying Lemma 3.2 in [8],

$$\lim_{\mu \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0, \theta \rightarrow 0} \mu \|x_{\mu, \varepsilon, \theta} - y_{\mu, \varepsilon, \theta}\|^2 = 0.$$

In (5.28), let  $\theta \downarrow 0$ , then  $\varepsilon \downarrow 0$ , then  $\mu \uparrow +\infty$ ; we obtain

$$0 < \delta_0/2 \leq 0.$$

A contradiction.  $\square$

## APPENDIX

### A.1. Verifying semigroup generation condition.

A.1.1. *The Cahn–Hilliard equation.* Let  $\omega \equiv \sup_r |V''(r)|^2/4$  and let  $C, C_n$  be defined according to (4.4) and (1.23). We prove the following.

LEMMA A.1. *The closure of  $C$  (resp.  $C_n$ ) is an  $m$ -dissipative operator in  $L^2(\mathcal{O})$ .*

The proof is divided into two parts.

LEMMA A.2. *Both  $C$  and  $C_n$  are dissipative.*

PROOF. Let  $x, y \in \mathcal{D}(C)$ ; then

$$\begin{aligned} & \langle Cx - Cy, x - y \rangle \\ &= \langle \Delta(-\Delta(x - y)), x - y \rangle + \langle V'(x) - V'(y), \Delta(x - y) \rangle - \omega \|x - y\|^2 \\ &\leq -\|\Delta(x - y)\|^2 + \sup_r |V''(r)| \|x - y\| \|\Delta(x - y)\| \\ &\quad - \frac{\sup_r |V''(r)|^2}{4} \|x - y\|^2 \leq 0. \end{aligned}$$

The case of  $C_n$  can be treated similarly.  $\square$

LEMMA A.3. *Let  $0 < \alpha \leq \alpha_0$  where  $\alpha_0$  is some prefixed small number. Suppose  $y_0 \in L^2(\mathcal{O})$ . Then for each  $y_n \equiv P_n y_0$ , there exists  $x_n \in \mathcal{R}(P_n)$  such that*

$$(A.1) \quad (I - \alpha C_n)x_n = y_n.$$

Moreover,

$$(A.2) \quad \|x_n\| \leq \|y_n\| \quad \text{and} \quad \|C_n x_n\| \leq (2/\alpha) \|y_n\|,$$

$$(A.3) \quad \|\Delta x_n\| \leq c(1 + \|y_n\|),$$

where  $c$  is a constant depending on  $F$  and  $\alpha$ . Consequently, there exists  $x_0 \in \mathcal{D}(\bar{C})$  such that

$$(I - \alpha \bar{C})x_0 = y_0.$$

PROOF. For each  $n$  fixed and finite, by its definition (1.23),  $C_n$  is Lipschitz on  $\text{Range}(P_n)$ . By the fixed point theorem, the essentially finite-dimensional equation (A.1) has a solution when  $\alpha > 0$  is sufficiently small. Since  $C_n$  is dissipative on  $\text{Range}(P_n)$ , solution actually exists for all  $\alpha > 0$ . See Lemma 2.13 of [26].

$\|x_n\| \leq \|y_n\|$  follows from the dissipativity of  $C_n$ . It follows then that  $\alpha \|C_n x_n\| = \|x_n - y_n\| \leq 2\|y_n\|$ .  $\langle x_n - \alpha C_n x_n, x_n \rangle = \langle y_n, x_n \rangle$ . That is,

$$\|x_n\|^2 - \alpha(-\|\Delta x_n\|^2 + \langle V'(x_n), \Delta x_n \rangle - \omega \|x_n\|^2) = \langle y_n, x_n \rangle,$$

which implies

$$\alpha(\|\Delta x_n\|^2 + \omega\|x_n\|^2) \leq \|y_n\|\|x_n\| - \|x_n\|^2 + \alpha\|V'(x_n)\|\|\Delta x_n\|.$$

Since  $V'(r)$  grows at most linearly, we can find constant  $c > 0$ ,

$$\|V'(x_n)\| \leq c(1 + \|x_n\|).$$

Hence

$$\begin{aligned} \alpha\|\Delta x_n\|^2 &\leq \|y_n\|\|x_n\| + \alpha c(1 + \|x_n\|)\|\Delta x_n\| \\ &\leq \|y_n\|^2 + \alpha c(1 + \|y_n\|)\|\Delta x_n\|, \end{aligned}$$

implying (A.3).

For each  $x \in H^2(\mathcal{O})$ ,  $\|\Delta x\| < \infty$  and

$$(A.4) \quad \Delta V'(x) = V''(x)\Delta x + V'''(x)\nabla x \nabla x.$$

We note that the Sobolev embedding  $H^1(\mathcal{O}) \rightarrow L^4(\mathcal{O})$  holds for space dimensions  $d = 1, 2, 3$ . Such result can be found in [1]: the case of  $d = 3$  follows from Lemma 5.10, the case of  $d = 2$  from Corollary 5.13 and the case of  $d = 1$  follows from Corollary 5.16 of [1]. Therefore,

$$\begin{aligned} \|\Delta V'(x_n)\| &\leq \sup_r |V''(r)|\|\Delta x_n\|_{L^2(\mathcal{O})} + \sup_r |V'''(r)|\|\nabla x_n\|_{L^4(\mathcal{O})}^2 \\ &\leq C(1 + \|\Delta x_n\|_{L^2(\mathcal{O})}^2) \end{aligned}$$

for some constant  $C$  independent of the  $x_n$ 's. By (A.3),  $\sup_n \|\Delta V'(x_n)\| < \infty$ . Using this estimate and  $\sup_n \|C_n x_n\| < \infty$ , we obtain  $\sup_n \|\Delta^2 x_n\| < \infty$ .

The boundedness of  $\sup_n (\|\Delta x_n\| + \|\Delta^2 x_n\|)$  implies that  $\Delta x_n$  is relatively compact in  $L^2(\mathcal{O})$ . Similarly, the boundedness of  $\|\Delta x_n\|$  and  $\|x_n\|$  implies the relative compactness of  $\nabla x_n$  and  $x_n$ . Selecting a subsequence if necessary, we have  $x_n \rightarrow x_0$ ,  $\Delta x_n \rightarrow \Delta x_0$  and  $\nabla x_n \rightarrow \nabla x_0$  for some  $x_0 \in H^2(\mathcal{O})$ . By (A.4),  $\Delta V'(x_n) \rightarrow \Delta V'(x_0)$ . Therefore

$$\|C_n x_n - C x_n\| = \|(P_n - I)\Delta V'(x_n)\| \rightarrow 0.$$

Noting

$$\begin{aligned} \alpha C x_n &= \alpha C_n x_n + \alpha(C x_n - C_n x_n) = x_n - y_n + \alpha(C x_n - C_n x_n) \rightarrow x_0 - y_0, \\ (x_0, \alpha^{-1}(x_0 - y_0)) &\in \bar{C}. \text{ By (A.1),} \end{aligned}$$

$$y_0 = x_0 - \alpha \frac{x_0 - y_0}{\alpha} \in (I - \alpha \bar{C})x_0. \quad \square$$

**A.1.2. The quasilinear equation with viscosity.** We consider the  $C_n, C$  in (1.27) and (4.5). Using similar a priori estimate arguments as in the Cahn–Hilliard equation case, we can prove the following:

**LEMMA A.4.**  *$C_n$  and the closure of  $C$  are both  $m$ -dissipative operators in  $L^2(\mathcal{O})$ .*

A.1.3. *The Allen–Cahn equation.* Let  $C_n$  and  $C$  be defined according to (1.18) and (4.1). Again, using a priori estimate arguments similar to the Cahn–Hilliard case, we can prove that both  $C_n$  and  $C$  are  $m$ -dissipative operators in  $L^2(\mathcal{O})$ . Alternatively, this conclusion can also be established by invoking the classical perturbation theory in, for instance, Corollary 6.19(i) of [26].

**A.2. Exponential compact containment estimates.** We illustrate the use of a stochastic Lyapunov function technique to verify Condition 1.12. We consider Examples 1.2, 1.5 and 1.8.

A.2.1. *Stochastic Allen–Cahn equation.* Recall (1.17) in Example 1.2:

$$(A.5) \quad dX_n(t) = \Delta P_n X_n(t) dt - P_n V'(P_n X_n(t)) dt + \frac{1}{\sqrt{n}} B_n(X_n(t)) dW(t).$$

The goal of this subsection is to prove the following lemma.

LEMMA A.5. *Condition 1.12 holds for  $X_n$ .*

We introduce the *free energy* function

$$\mathcal{E}(x) \equiv \frac{1}{2} \|\nabla x\|^2 + \int_{\mathcal{O}} V(x) d\theta.$$

First, we prove the following estimate: for every  $T, a > 0$  and  $C_0 > 0$ , there exists  $C_1 > 0$  such that

$$(A.6) \quad \sup_{x: \mathcal{E}(x) \leq C_0} P(\mathcal{E}(X_n(t)) > C_1, \text{ some } 0 < t \leq T | X_n(0) = x) \leq e^{-na}.$$

Let us approximate  $\mathcal{E}$  by

$$(A.7) \quad \mathcal{E}_n(x) \equiv -\frac{1}{2} \langle \Delta P_n x, x \rangle + \int_{\mathcal{O}} V(P_n x(\theta)) d\theta.$$

Note that if  $X_n(0) \in \mathcal{R}(P_n)$ , the range of  $P_n$ , then  $X_n(t) \in \mathcal{R}(P_n)$ , hence  $\mathcal{E}_n(X_n(t)) = \mathcal{E}(X_n(t))$ . Define

$$(A.8) \quad f_n(x) \equiv \log \left( 1 + \frac{1}{M^2} \mathcal{E}_n(x) \right)$$

where  $M \equiv \sup_{\theta, x} |\sigma(\theta, x)| < \infty$ . Then

$$Df_n(x) = \frac{(-1/M^2)(\Delta P_n x - P_n V'(P_n x))}{1 + (1/M^2)\mathcal{E}_n(x)}$$

and

$$D^2 f_n(x) = \frac{(-1/M^2)(\Delta P_n x - P_n V'(P_n x)) \otimes (1/M^2)(\Delta P_n x - P_n V'(P_n x))}{(1 + (1/M^2)\mathcal{E}_n(x))^2} \\ + \frac{(1/M^2)(-\Delta P_n + P_n V''(P_n x))}{1 + (1/M^2)\mathcal{E}_n(x)},$$

where  $P_n V''(P_n x)$  means a linear operator on  $L^2(\mathcal{O})$  [for each  $x \in L^2(\mathcal{O})$  fixed]:

$$(P_n V''(P_n x))y \equiv \sum_{k \equiv (k_1, \dots, k_d) = (1, \dots, 1)}^{(m_n, \dots, m_n)} \langle V''(P_n x)y, e_k \rangle e_k \quad \forall y \in L^2(\mathcal{O}).$$

Then

$$\begin{aligned} H_n f_n(x) &\equiv \langle \Delta P_n x - P_n V'(P_n x), Df_n(x) \rangle \\ &\quad + \frac{1}{2} \|B_n^*(x) Df_n(x)\|_{U_0}^2 + \frac{1}{2n} \text{Tr}(D^2 f_n(x) B_n(x) B_n^*(x)) \\ &= \frac{(-1/M^2) \|\Delta P_n x - P_n V'(P_n x)\|^2}{1 + (1/M^2)\mathcal{E}_n(x)} \\ &\quad + \frac{1}{2} \left(1 - \frac{1}{n}\right) \left\| \frac{(1/M^2) B_n^*(x) (\Delta P_n x - P_n V'(P_n x))}{1 + (1/M^2)\mathcal{E}_n(x)} \right\|^2 \\ &\quad + \frac{1}{2n} \left( \sum_{k=(1, \dots, 1)}^{(m_n, \dots, m_n)} \left\langle (-\Delta P_n + P_n V''(P_n x)) \frac{B_n(x)}{M^2} e_k, \frac{B_n(x)}{M^2} e_k \right\rangle \right) \\ &\quad \times \left(1 + \frac{1}{M^2} \mathcal{E}_n(x)\right)^{-1} \\ &= \frac{(-1/M^2) \|\Delta P_n x - P_n V'(P_n x)\|^2}{1 + (1/M^2)\mathcal{E}_n(x)} \\ &\quad + \frac{1}{2} \left(1 - \frac{1}{n}\right) \left\| \frac{(1/M^2) B_n^*(x) (\Delta P_n x - P_n V'(P_n x))}{1 + (1/M^2)\mathcal{E}_n(x)} \right\|^2 \\ (A.9) \quad &\quad + \frac{1}{2n} \frac{\sum_{k=(1, \dots, 1)}^{(m_n, \dots, m_n)} \sum_{i \equiv (i_1, \dots, i_d) = (1, \dots, 1)}^{(m_n, \dots, m_n)} \lambda_i(\langle (B_n(x)/M^2) e_k, e_i \rangle)^2}{1 + (1/M^2)\mathcal{E}_n(x)} \\ &\quad + \frac{1}{2n} \left( \sum_{k=(1, \dots, 1)}^{(m_n, \dots, m_n)} \sum_{i=(1, \dots, 1)}^{(m_n, \dots, m_n)} \sum_{j=(1, \dots, 1)}^{(m_n, \dots, m_n)} \left\langle \frac{\sigma(\cdot, x)}{M^2} e_k, e_j \right\rangle \langle V''(P_n x) e_j, e_i \rangle \right. \\ &\quad \left. \times \left\langle \frac{\sigma(\cdot, x)}{M^2} e_k, e_i \right\rangle \right) \end{aligned}$$



$$\begin{aligned}
& \times \left(1 + \frac{1}{M^2} \mathcal{E}_n(x)\right)^{-1} \\
& \leq \left\{ -\frac{1}{1 + (1/M^2) \mathcal{E}_n(x)} + \frac{1}{2} \left(1 - \frac{1}{n}\right) \left(\frac{1}{1 + (1/M^2) \mathcal{E}_n(x)}\right)^2 \right\} \\
& \quad \times \frac{\|\Delta P_n x - P_n V'(P_n x)\|^2}{M^2} \\
& \quad + \frac{1}{2n} m_n^d \sum_{i=(1, \dots, 1)}^{(m_n, \dots, m_n)} \lambda_i + \frac{1}{n} m_n^{3d} \int_{\mathcal{O}} |V''(P_n x(\theta))| d\theta \\
& \leq 0 + \frac{4^d \pi^{2d} (1 + m_n)^{4d}}{6n} + \frac{m_n^{3d}}{n} \sup_r |V''(r)| \leq \text{Constant} < \infty,
\end{aligned}$$

where the  $\lambda_i$ 's are eigenvalues defined in (1.12) and the constant is independent of  $n$ . In the last inequality above, we used (1.14), and the estimate that

$$(A.10) \quad \sum_{i=(1, \dots, 1)}^{(m_n, \dots, m_n)} \lambda_i = \left( \sum_{i_1=1}^{m_n} \mu_{i_1} \right)^d \leq 4^d \pi^{2d} \frac{(1 + m_n)^{3d}}{3},$$

where  $\mu_i$  is the one in (1.10).

Let

$$\tau_n \equiv \inf\{t > 0 : \mathcal{E}_n(X_n(t)) \geq C_1\}.$$

By optional sampling theorem

$$\begin{aligned}
& \sup_{x: \mathcal{E}(x) \leq C_0} P(\mathcal{E}_n(X_n(t)) > C_1, \text{ some } 0 < t \leq T | X_n(0) = x) \\
& \quad \times e^{n(C_1 - C_0) - nT \sup_{n,x} H_n f_n(x)} \\
(A.11) \quad & \leq \sup_{x: \mathcal{E}(x) \leq C_0} E[e^{nf_n(X_n(T \wedge \tau_n)) - nf_n(X_n(0)) - \int_0^{T \wedge \tau_n} n H_n f_n(X_n(s)) ds} | X_n(0) = x] \\
& = 1.
\end{aligned}$$

Hence (A.6) follows. We now relax the initial condition in the estimate to that in Lemma A.5/Condition 1.12. We achieve this by the following result:

LEMMA A.6. *We denote by  $X_n^x$  the solution of (A.5) with initial value  $X_n(0) = x$ . Then for each  $T, a > 0$ , there exists a constant  $C = C(T, a) > 0$  such that*

$$\begin{aligned}
P\left(\sup_{0 \leq t \leq T} \|X_n^x(t) - X_n^y(t)\| > C\varepsilon \mid \|x - y\| < \varepsilon\right) & < e^{-na} \\
\forall 0 < \varepsilon < 1, n = 1, 2, \dots
\end{aligned}$$

PROOF. For each  $n$  fixed,  $(X_n^x(t), X_n^y(t))$  is a two-component Markov process that solves the martingale problem with generator

$$\begin{aligned}\mathcal{A}_n f(x, y) &= \langle \Delta P_n x - P_n V'(P_n x), D_x f(x, y) \rangle \\ &\quad + \langle \Delta P_n y - P_n V'(P_n y), D_y f(x, y) \rangle \\ &\quad + \frac{1}{2n} \text{Tr}((D_{xx}^2 f) B_n(x) B_n^*(x)) \\ &\quad + (D_{yy}^2 f) B_n(y) B_n^*(y) + 2(D_{xy}^2 f) B_n(x) B_n^*(y)\end{aligned}$$

for  $f(x, y) \in C^2(L^2(\mathcal{O}) \times L^2(\mathcal{O}))$ . Let  $\varepsilon > 0$  and

$$f_{n,\varepsilon}(x, y) \equiv \log \left( 1 + \frac{1}{2} \left\| \frac{P_n x - P_n y}{\varepsilon} \right\|^2 \right).$$

It follows that

$$\mathcal{H}_n f_{n,\varepsilon}(x, y) \equiv \frac{1}{n} e^{-n f_{n,\varepsilon}} \mathcal{A}_n e^{n f_{n,\varepsilon}}(x, y) \leq C_0 < \infty,$$

where constant  $C_0$  is independent of  $n$  as well as  $\varepsilon$ . By an argument identical to that used in the proof of (A.11), the conclusion follows.  $\square$

PROOF OF LEMMA A.5. Let compact set  $K \subset L^2(\mathcal{O})$  and  $a, T, \varepsilon > 0$ . It is enough for us to show that for any  $x_n \in K$ , there exists compact set  $K_1 \subset E$ ,

$$P(\exists t \in [0, T], X_n^{x_n}(t) \notin K_1^{2\varepsilon}) \leq 2e^{-na}.$$

Let  $\delta > 0$ . By compactness of  $K$ , there exists  $\{x_0^1, \dots, x_0^{m(\delta)}\}$  such that

$$K \subset \bigcup_{k=1}^{m(\delta)} B(x_0^k, \delta).$$

Since  $\{x : \mathcal{E}(x) < +\infty\} \subset H^1(\mathcal{O})$  is dense in  $L^2(\mathcal{O})$ , we can select  $x_0^k$  so that

$$\sup_{j=1, \dots, m(\delta)} \mathcal{E}(x_0^j) < +\infty.$$

Therefore, we can choose

$$x_{0,n} \in \{x_0^{(1)}, \dots, x_0^{(m(\delta))}\}$$

such that  $\|x_{0,n} - x_n\| < \delta$ . By Lemma A.6, there exists  $C = C(T, a) > 0$  (independent of  $\delta$ ) such that

$$P\left(\sup_{0 \leq t \leq T} \|X_n^{x_{0,n}}(t) - X_n^{x_n}(t)\| > C\delta\right) < e^{-na}.$$

By (A.6) and the compactness of level sets for  $\mathcal{E}$ , there exists a compact set  $K_1 \subset E$  such that

$$P(X_n^{x_0,n}(t) \notin K_1^\delta, \exists t \in [0, T]) \leq e^{-na}.$$

It follows that

$$\begin{aligned} & \{\exists t \in [0, T], X_n^{x_n}(t) \notin K_1^{(1+C)\delta}\} \\ & \subset \{\exists t \in [0, T], X_n^{x_0,n}(t) \notin K_1^\delta\} \cup \left\{ \sup_{0 \leq t \leq T} \|X_n^{x_0,n}(t) - X_n^{x_n}(t)\| > C\delta \right\}. \end{aligned}$$

Therefore

$$\begin{aligned} & P(\exists t \in [0, T], X_n^{x_n}(t) \notin K_1^{(1+C)\delta}) \\ & \leq P(\exists t \in [0, T], X_n^{x_0,n}(t) \notin K_1^\delta) + P\left(\sup_{0 \leq t \leq T} \|X_n^{x_0,n}(t) - X_n^{x_n}(t)\| > C\delta\right) \\ & \leq 2e^{-na}. \end{aligned}$$

Taking  $\delta = \varepsilon/(1+C)$ , we complete the proof.  $\square$

**A.2.2. Stochastic Cahn–Hilliard equation.** Recall the stochastic Cahn–Hilliard equation (1.22) in Example 1.5:

$$(A.12) \quad dX_n(t) = \Delta P_n(-\Delta P_n X_n(t) + P_n V'(P_n X_n(t))) dt + \frac{1}{\sqrt{n}} B_n dW(t).$$

We prove the following lemma.

LEMMA A.7. *Condition 1.12 holds for  $X_n$ .*

Using identical arguments as in the Allen–Cahn case, we just need the following estimates:  $\sup_n \sup_x H_n f_n(x) < \infty$  for the  $f_n$  below, and (A.15).

Define  $\mathcal{E}, \mathcal{E}_n$  the same way as in (A.7). Let

$$f_n(x) \equiv \log\left(1 + \frac{1}{M^2} \mathcal{E}_n(x)\right),$$

where  $M > 0$  is the constant in the Poincaré type inequality

$$(A.13) \quad \left\| x - \int_{\theta \in \mathcal{O}} x(\theta) d\theta \right\| \leq M \|\nabla x\| \quad \forall x \in H^2(\mathcal{O}).$$

Let  $\lambda_k$  be defined according to (1.12). Then

$$\begin{aligned} H_n f_n(x) & \equiv \langle \Delta P_n(-\Delta P_n x + P_n V'(P_n x)), Df_n(x) \rangle \\ & + \frac{1}{2} \|Df_n(x)\|^2 + \frac{1}{2n} \text{Tr}(D^2 f_n(x)) \end{aligned}$$

$$\begin{aligned}
&= \frac{(-1/M^2) \|(-\Delta P_n)^{1/2} (\Delta P_n x - P_n V'(P_n x))\|^2}{1 + (1/M^2) \mathcal{E}_n(x)} \\
&\quad + \frac{1}{2} \left(1 - \frac{1}{n}\right) \frac{\|\Delta P_n x - P_n V'(P_n x)\|^2 / M^4}{(1 + (1/M^2) \mathcal{E}_n(x))^2} \\
&\quad + \frac{1}{2n} \frac{(1/M^2) \sum_{k=(1,\dots,1)}^{(m_n,\dots,m_n)} \langle (-\Delta P_n + P_n V''(P_n x)) e_k, e_k \rangle}{1 + (1/M^2) \mathcal{E}_n(x)} \\
&= \frac{(-1/M^2) \|\nabla(\Delta P_n x - P_n V'(P_n x))\|^2}{1 + (1/M^2) \mathcal{E}_n(x)} \\
&\quad + \frac{1}{2} \left(1 - \frac{1}{n}\right) \frac{(1/M^2) \|\Delta P_n x - P_n V'(P_n x)\|^2 / M^2}{(1 + (1/M^2) \mathcal{E}_n(x))^2} \\
&\quad + \frac{1}{M^2} \frac{1}{2n} \frac{\sum_{k=(1,\dots,1)}^{(m_n,\dots,m_n)} \lambda_k + \sum_{k=(1,\dots,1)}^{(m_n,\dots,m_n)} \langle V''(P_n x) e_k, e_k \rangle}{1 + (1/M^2) \mathcal{E}_n(x)} \\
(A.14) \quad &\leq \left\{ -\frac{1}{1 + (1/M^2) \mathcal{E}_n(x)} + \frac{1}{2} \left(1 - \frac{1}{n}\right) \left( \frac{1}{1 + (1/M^2) \mathcal{E}_n(x)} \right)^2 \right\} \\
&\quad \times \frac{\|\nabla(\Delta P_n x - P_n V'(P_n x))\|^2}{M^2} \\
&\quad + \frac{1}{2} \left( \frac{1}{M^2} \int_{\mathcal{O}} V'(P_n x(\theta)) d\theta \right)^2 \\
&\quad + \frac{1}{2n} \frac{1}{M^2} \sum_{k=(1,\dots,1)}^{(m_n,\dots,m_n)} \left( \lambda_k + \sup_r |V''(r)| \right) \\
&\leq C < \infty.
\end{aligned}$$

In the above derivations, we used (A.13):

$$\begin{aligned}
&\|\Delta P_n x - P_n V'(P_n x)\|^2 \\
&= \left\{ \int_{\mathcal{O}} (\Delta P_n x(\theta) - P_n V'(P_n x(\theta))) d\theta \right\}^2 \\
&\quad + \left\| \Delta P_n x - P_n V'(P_n x) - \int_{\mathcal{O}} (\Delta P_n x(\theta) - P_n V'(P_n x(\theta))) d\theta \right\|^2 \\
&\leq \left( \int_{\mathcal{O}} V'(P_n x(\theta)) d\theta \right)^2 + M^2 \|\nabla(\Delta P_n x - P_n V'(P_n x))\|^2.
\end{aligned}$$

We also made use of (A.10) and condition (1.21).

LEMMA A.8. *We denote by  $X_n^x$  the solution of (A.12) with initial value  $X_n(0) = x$ . Then for each  $T, a > 0$ , there exists a constant  $C = C(T, a) > 0$*

such that

$$(A.15) \quad P\left(\sup_{0 \leq t \leq T} \|X_n^x(t) - X_n^y(t)\| > C\varepsilon \mid \|x - y\| < \varepsilon\right) < e^{-na} \\ \forall 0 < \varepsilon < 1, n = 1, 2, \dots$$

PROOF. The proof follows the same idea as in Lemma A.6.  $\square$

A.2.3. *Stochastic quasilinear equation with viscosity.* Using the same ideas as in the stochastic Allen–Cahn and Cahn–Hilliard case, by choosing

$$\mathcal{E}(x) = \frac{1}{2}(\|x\|^2 + \|\nabla x\|^2), \quad \mathcal{E}_n(x) = \frac{1}{2}(\|P_n x\|^2 - \langle \Delta P_n x, x \rangle)$$

and

$$f_n(x) = \log(1 + \alpha \mathcal{E}_n(x)),$$

we can prove the following.

LEMMA A.9. Condition 1.12 holds for the  $X_n$  in (1.26).

**A.3. Approximations of the Tataru distance function.** Let  $E, U_0$  be real separable Hilbert spaces. We discuss approximations of the *Tataru distance function*  $d_C$  (Definition 1.13) by  $C^2(E)$  functions. Throughout this section, we assume Condition 1.11 is satisfied for  $C, C_n$ .

We want to keep two useful properties (3.3) and (3.4) in the approximation. The functions  $h_{\varepsilon, y}(x)$  and  $h_{n, \varepsilon, y}(x)$  defined in (3.8) and (3.9) satisfy these requirements—see (A.18) and (A.20).

LEMMA A.10. For each  $\varepsilon > 0$  small enough, define  $\phi_\varepsilon$  according to (3.7):  $\phi_\varepsilon(r) = \sqrt{r}$  when  $r \geq \varepsilon$  and

$$\phi_\varepsilon(r) = \sqrt{\varepsilon} + \frac{r - \varepsilon}{2\sqrt{\varepsilon}} - \frac{(r - \varepsilon)^2}{8\varepsilon\sqrt{\varepsilon}} \quad \text{when } 0 \leq r \leq \varepsilon.$$

Then:

- (1)  $\phi'_\varepsilon, \phi''_\varepsilon \in C_b([0, +\infty))$ ;  $\phi_\varepsilon$  is nondecreasing,  $\sup_r r |\phi''_\varepsilon(r)| < +\infty$ .
- (2)  $\lim_{\varepsilon \rightarrow 0} \sup_{r \geq 0} |\phi_\varepsilon(r) - \sqrt{r}| = 0$ .
- (3)

$$(A.16) \quad 0 \leq r \phi'_\varepsilon(r^2) \leq 1/2.$$

PROOF. The first two properties follow from direct verification. To see that the third one holds, let  $f(r) = r \phi'_\varepsilon(r^2)$ ; then  $f'(r) > 0$  for  $0 < r < \varepsilon$ , and  $f'(r) = 0$  when  $r \geq \varepsilon$ . Hence  $r \phi'_\varepsilon(r^2) \leq \varepsilon \phi'_\varepsilon(\varepsilon^2) = 1/2$ .  $\square$

LEMMA A.11. *Let  $\varphi, \varphi_n : [0, \infty) \rightarrow [0, \infty)$  be continuous. Suppose there exist  $0 < m < M < +\infty$ ,  $0 \leq c < \infty$  such that  $mt \leq \varphi_n(t) \leq c + Mt$ ,  $n = 1, 2, \dots$ . Suppose further that  $0 < a_n \rightarrow +\infty$  and that*

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} |\varphi_n(t) - \varphi(t)| = 0 \quad \forall T \geq 0.$$

Then

$$\lim_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_0^\infty e^{-a_n \varphi_n(t)} dt = \inf_{t \geq 0} \varphi(t).$$

PROOF. It is straightforward to see that, when  $T > 0$  is large enough but fixed,

$$\lim_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_0^T e^{-a_n \varphi_n(t)} dt = \inf_{0 \leq t \leq T} \varphi(t) = \inf_{t \geq 0} \varphi(t).$$

Take  $T \geq \varphi(0)/m$ ; then when  $n$  is large enough

$$-\frac{1}{a_n} \log \int_T^\infty e^{-a_n \varphi_n(t)} dt \geq -\frac{1}{a_n} \log \int_T^\infty e^{-a_n m t} dt \geq mT \geq \varphi(0) \geq \inf_{t \geq 0} \varphi(t).$$

Therefore,

$$\begin{aligned} \inf_{t \geq 0} \varphi(t) &= \min \left\{ \liminf_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_0^T e^{-a_n \varphi_n(t)} dt, \liminf_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_T^\infty e^{-a_n \varphi_n(t)} dt \right\} \\ &\leq \liminf_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_0^\infty e^{-a_n \varphi_n(t)} dt \leq \limsup_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_0^\infty e^{-a_n \varphi_n(t)} dt \\ &\leq \limsup_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_0^T e^{-a_n \varphi_n(t)} dt = \inf_{t \geq 0} \varphi(t). \quad \square \end{aligned}$$

LEMMA A.12. *Let  $a_n > 0$  be such that  $\lim_{n \rightarrow \infty} a_n = \infty$ . Define  $\phi_\varepsilon$  as in (3.7) and  $h_\varepsilon, h_{n,\varepsilon}$  according to (3.8) and (3.9):*

$$h_{n,\varepsilon}(x) \equiv -\frac{1}{a_n} \log \int_0^\infty e^{-a_n \{t + \phi_\varepsilon(\|x - S_n(t)y\|^2)\}} dt$$

and

$$h_\varepsilon(x) \equiv \inf_{t \geq 0} \{t + \phi_\varepsilon(\|x - S(t)y\|^2)\}.$$

Then:

- (1)  $h_{n,\varepsilon}(x) \geq c$  whenever  $\|x\| \geq \|y\| + c$ , for every  $c > 1$ ,  $a_n > 1$ .
- (2) For each  $y \in E$  fixed,

$$(A.17) \quad \lim_{\varepsilon \rightarrow 0^+} \sup_{x \in E} |h_\varepsilon(x) - d_C(x, y)| = 0.$$

(3) For each  $\varepsilon > 0$  fixed,

$$\lim_{n \rightarrow \infty} \sup_{x \in K} |h_\varepsilon(x) - h_{n,\varepsilon}(x)| = 0 \quad \forall \text{ compact } K \subset E.$$

(4)  $h_{n,\varepsilon} \in C^2(E)$ ;

$$(A.18) \quad \|Dh_{n,\varepsilon}(x)\| \leq 1.$$

If  $B_n$  is a Hilbert-Schmidt operator from  $U_0$  to  $E$  [i.e.,  $B_n \in L_2(U_0, E)$ ], then

$$(A.19) \quad \begin{aligned} & |\operatorname{Tr}[D^2 h_{n,\varepsilon}(x) B_n B_n^*]| \\ & \leq \left( 2a_n + 4 \sup_{r \geq 0} r |\phi_\varepsilon''(r)| + 2 \sup_{r \geq 0} \phi_\varepsilon'(r) \right) \|B_n\|_{L_2(U_0, E)}^2. \end{aligned}$$

(5) Let  $\varphi \in \mathcal{T}$ ; we have

$$(A.20) \quad \limsup_{r \rightarrow 0+, z \rightarrow x} \frac{\varphi(h_{n,\varepsilon}(S_n(r)z)) - \varphi(h_{n,\varepsilon}(z))}{r} \leq \sup_{r \geq 0} \varphi'(r) \quad \forall x \in E.$$

If  $B_n \in L_2(U_0, E)$ , then

$$(A.21) \quad \begin{aligned} & |\operatorname{Tr}[D^2(\varphi \circ h_{n,\varepsilon})(x) B_n B_n^*]| \\ & \leq \left\{ \sup_{s \geq 0} \varphi''(s) + \sup_{s \geq 0} \varphi'(s) \left( 2a_n + 4 \sup_{r \geq 0} r |\phi_\varepsilon''(r)| + 2 \sup_{r \geq 0} \phi_\varepsilon'(r) \right) \right\} \\ & \quad \times \|B_n\|_{L_2(U_0, E)}^2. \end{aligned}$$

PROOF. Part (1): Since  $\|S_n(t)y\| \leq \|y\|$ ,  $(\|x\| - \|y\|) \vee 0 \leq \|x - S_n(t)y\|$ . Hence

$$\phi_\varepsilon((\|x\| - \|y\|) \vee 0)^2 \leq h_{n,\varepsilon}(x)$$

when  $a_n > 1$ . Noting  $\phi_\varepsilon(r^2) = r$  when  $r \geq 1$ , the conclusion follows.

Part (2) is a direct consequence of part (2) of Lemma A.10.

Part (3) follows if we prove that for each  $x_n \rightarrow x$ ,

$$\lim_{n \rightarrow \infty} h_{n,\varepsilon}(x_n) = h_\varepsilon(x).$$

Take  $\varphi_n(t) = \{t + \phi_\varepsilon(\|x_n - S_n(t)y\|^2)\}$ ,  $\varphi(t) = \{t + \phi_\varepsilon(\|x - S(t)y\|^2)\}$ . For each  $T > 0$ ,

$$\begin{aligned} \sup_{0 \leq t \leq T} |\varphi_n(t) - \varphi(t)| &= \sup_{0 \leq t \leq T} |\phi_\varepsilon(\|x_n - S_n(t)y\|^2) - \phi_\varepsilon(\|x - S(t)y\|^2)| \\ &\leq \sup_{r \geq 0} \phi_\varepsilon'(r) \sup_{0 \leq t \leq T} \left| \|x_n - S_n(t)y\| - \|x - S(t)y\| \right| \rightarrow 0 \end{aligned}$$

as  $n \rightarrow \infty$ .

In addition

$$t \leq \varphi_n(t) \leq t + \sup_n \phi_\varepsilon((\|x_n\| + \|y\|)^2).$$

Apply Lemma A.11; therefore

$$\lim_{n \rightarrow \infty} -\frac{1}{a_n} \log \int_0^\infty e^{-a_n \varphi_n(t)} dt = \inf_{t \geq 0} \varphi(t).$$

Part (4): It can be verified that

$$Dh_{n,\varepsilon}(x) = 2 \frac{\int_0^\infty e^{-a_n \{t + \phi_\varepsilon(\|x - S_n(t)y\|^2)\}} \phi'_\varepsilon(\|x - S_n(t)y\|^2) (x - S_n(t)y) dt}{\int_0^\infty e^{-a_n \{t + \phi_\varepsilon(\|x - S_n(t)y\|^2)\}} dt} \in E.$$

By Lemma A.10,

$$\|Dh_{n,\varepsilon}(x)\| \leq 1.$$

Direct calculation also gives

$$\begin{aligned} D^2 h_{n,\varepsilon}(x) &= \frac{1}{\int_0^\infty e^{-a_n \{t + \phi_\varepsilon(\|x - S_n(t)y\|^2)\}} dt} \\ &\quad \times \int_0^\infty e^{-a_n \{t + \phi_\varepsilon(\|x - S_n(t)y\|^2)\}} \\ &\quad \times ((-4a_n)(\phi'_\varepsilon(\|x - S_n(t)y\|^2))^2 \\ &\quad \times (x - S_n(t)y) \otimes (x - S_n(t)y) \\ &\quad + 4\phi''_\varepsilon(\|x - S_n(t)y\|^2)(x - S_n(t)y) \otimes (x - S_n(t)y) \\ &\quad + 2\phi'_\varepsilon(\|x - S_n(t)y\|^2)I) dt \\ &\quad + a_n Dh_{n,\varepsilon}(x) \otimes Dh_{n,\varepsilon}(x). \end{aligned}$$

Let  $\{\hat{e}_1, \dots, \hat{e}_k, \dots\}$  be a complete orthonormal basis for  $E$ . Then,

$$\begin{aligned} &|\text{Tr}[D^2 h_{n,\varepsilon}(x) B_n B_n^*]| \\ &= \left| \sum_k \langle D^2 h_{n,\varepsilon}(x) B_n B_n^* \hat{e}_k, \hat{e}_k \rangle \right| \\ &\leq \sup_{t \geq 0} (4a_n (\phi'_\varepsilon(\|x - S_n(t)y\|^2))^2 + 4|\phi''_\varepsilon|(\|x - S_n(t)y\|^2)) \\ &\quad \times \|B_n^*(x - S_n(t)y)\|_{U_0}^2 \\ &\quad + 2 \sup_{t \geq 0} \phi'_\varepsilon(\|x - S_n(t)y\|^2) \|B_n\|_{L_2(U_0, E)}^2 + a_n \|B_n^* Dh_{n,\varepsilon}(x)\|^2 \\ &\leq 4a_n \sup_{t \geq 0} (\|x - S_n(t)y\| \phi'_\varepsilon(\|x - S_n(t)y\|^2) \|B_n\|_{L_2(U_0, E)})^2 \end{aligned}$$



$$\begin{aligned}
& + 4 \sup_{t \geq 0} |\phi''_\varepsilon|(\|x - S_n(t)y\|^2) \|x - S_n(t)y\|^2 \|B_n\|_{L_2(U_0, E)}^2 \\
& + 2 \sup_{t \geq 0} \phi'_\varepsilon(\|x - S_n(t)y\|^2) \|B_n\|_{L_2(U_0, E)}^2 + a_n \|Dh_{n, \varepsilon}(x)\|^2 \|B_n\|_{L_2(U_0, E)}^2 \\
& \leq \left( 4a_n \sup_{r \geq 0} (r \phi'_\varepsilon(r^2))^2 + 4 \sup_{r \geq 0} r |\phi''_\varepsilon(r)| + 2 \sup_{r \geq 0} \phi'_\varepsilon(r) + a_n \right) \|B_n\|_{L_2(U_0, E)}^2.
\end{aligned}$$

To derive the second inequality above, we used the following: let  $e_1, \dots, e_k, \dots$  be a complete orthonormal system for  $U_0$ ; then for each  $z \in E$ ,

$$\begin{aligned}
\|B_n^* z\|^2 &= \sum_k \langle B_n^* z, e_k \rangle^2 = \sum_k \langle z, B_n e_k \rangle^2 \leq \sum_k \|z\|^2 \|B_n e_k\|^2 \\
&= \|z\|^2 \|B_n\|_{L_2(U_0, E)}^2.
\end{aligned}$$

Noting (A.16), (A.19) holds.

Part (5): Let  $r > 0$ ; then

$$\begin{aligned}
h_{n, \varepsilon}(x) + r &= -\frac{1}{a_n} \log \int_0^\infty e^{-a_n(t+r+\phi_\varepsilon(\|x-S_n(t)y\|^2))} dt \\
&\geq -\frac{1}{a_n} \log \int_0^\infty e^{-a_n(t+r+\phi_\varepsilon(\|S_n(r)x-S_n(t+r)y\|^2))} dt \\
&= -\frac{1}{a_n} \log \int_r^\infty e^{-a_n(t+\phi_\varepsilon(\|S_n(r)x-S_n(t)y\|^2))} dt \\
&\geq h_{n, \varepsilon}(S_n(r)x).
\end{aligned}$$

Hence we have (A.20). Equation (A.21) follows from (A.18), (A.19) and

$$D^2\varphi \circ h_{n, \varepsilon}(x) = \varphi'' \circ h_{n, \varepsilon}(x)(Dh_{n, \varepsilon}(x) \otimes Dh_{n, \varepsilon}(x)) + \varphi' \circ h_{n, \varepsilon}(x) D^2 h_{n, \varepsilon}(x).$$

□

**Acknowledgments.** I would like to thank Professor Tom Kurtz for useful discussions on stochastic equation in infinite dimensions, and Professor Markos Katsoulakis for useful discussions on stochastic Allen–Cahn and Cahn–Hilliard equations. An anonymous referee carefully read through the paper, and provided an extensive list of thoughtful comments and suggestions. These are incorporated in the paper; I gratefully acknowledge the help as well.

## REFERENCES

- [1] ADAMS, R. (1975). *Sobolev Spaces*. Academic Press, New York. [MR450957](#)
- [2] BARLES, G. and PERTHAME, B. (1988). Exit time problems in optimal control and vanishing viscosity solution method. *SIAM J. Control Optim.* **26** 1133–1148. [MR957658](#)

- [3] BARLES, G. and PERTHAME, B. (1990). Comparison results in Dirichlet type first order Hamilton–Jacobi equations. *Appl. Math. Optim.* **21** 21–44. [MR1014943](#)
- [4] CHOW, P.-L. (1992). Large deviation problem for some parabolic Itô equations. *Comm. Pure Appl. Math.* **45** 97–120. [MR1135925](#)
- [5] CRANDALL, M. G., ISHII, H. and LIONS, P.-L. (1992). User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N. S.)* **27** 1–67. [MR1118699](#)
- [6] CRANDALL, M. G. and LIGGETT, T. M. (1971). Generation of semigroups of non-linear transformations on general Banach spaces. *Amer. J. Math.* **93** 265–298. [MR287357](#)
- [7] CRANDALL, M. G. and LIONS, P.-L. (1985). Hamilton–Jacobi equations in infinite dimensions, Part I. *J. Funct. Anal.* **62** 379–396; Part II. *ibid.* **65** (1986) 368–405; Part III. *ibid.* **68** (1986) 214–247; Part IV. *ibid.* **90** (1990) 237–283; Part V. *ibid.* **97** (1991) 417–465; Part VII. *ibid.* **125** (1994) 111–148.
- [8] CRANDALL, M. G. and LIONS, P.-L. (1994). Hamilton–Jacobi equations in infinite dimensions, Part VI: Nonlinear  $A$  and Tataru’s method refined. *Lecture Notes in Pure and Appl. Math.* **155** 51–89. Dekker, New York. [MR1254890](#)
- [9] DA PRATO, G. and ZABCZYK, J. (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge Univ. Press. [MR1207136](#)
- [10] DAWSON, D. and GÄRTNER, J. (1987). Large deviation from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308. [MR885876](#)
- [11] DE ACOSTA, A. (1997). Exponential tightness and projective systems in large deviation theory. In *Festschrift for Lucien Le Cam* 143–156. Springer, New York. [MR1462943](#)
- [12] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. Springer, New York. [MR1619036](#)
- [13] DONSKER, M. D. and VARADHAN, S. R. S. (1975). Asymptotic evaluation of certain Markov process expectations for large time, I. *Comm. Pure Appl. Math.* **28** 1–47; II *ibid.* **28** (1975) 279–301; III *ibid.* **29** (1976) 389–461.
- [14] EKELAND, I. (1979). Nonconvex minimization problems. *Bull. Amer. Math. Soc. (N. S.)* **1** 443–474. [MR526967](#)
- [15] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes*. Wiley, New York. [MR838085](#)
- [16] EVANS, L. C. and ISHII, H. (1985). A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2** 1–20. [MR781589](#)
- [17] FENG, J. (1999). Martingale problems for large deviations of Markov processes. *Stoch. Process. Appl.* **81** 165–216. [MR1694569](#)
- [18] FENG, J. and KURTZ, T. G. (2003). Large deviations of stochastic processes. Preprint.
- [19] FLEMING, W. H. (1978). Exit probabilities and optimal stochastic control. *Appl. Math. Optim.* **4** 329–346. [MR512217](#)
- [20] FLEMING, W. H. and SONER, H. M. (1993). *Controlled Markov Processes and Viscosity Solutions*. Springer, New York.
- [21] FLEMING, W. H. and SOUGANIDIS, P. E. (1986). Asymptotic series and the method of vanishing viscosity. *Indiana Univ. Math. J.* **35** 425–447. [MR833404](#)
- [22] FLEMING, W. H. and SOUGANIDIS, P. E. (1986). A PDE approach to asymptotic estimates for optimal exit probabilities. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **13** 171–192. [MR876121](#)
- [23] FREIDLIN, M. I. and WENTZELL, A. D. (1983). *Random Perturbations of Dynamical Systems*. Springer, New York. [MR1652127](#)

- [24] KALLIANPUR, G. and XIONG, J. (1996). Large deviations for a class of stochastic partial differential equations. *Ann. Probab.* **24** 320–345. [MR1387638](#)
- [25] KURTZ, T. G. (1987). Martingale problems for controlled processes. *Lecture Notes in Control and Inform. Sci.* **91** 75–90. Springer, Berlin. [MR894107](#)
- [26] MIYADERA, I. (1992). *Nonlinear Semigroups*. Amer. Math. Soc., Providence, RI.
- [27] PESZAT, S. (1994). Large deviation principle for stochastic evolution equations. *Probab. Theory Related Fields* **98** 113–136. [MR1254827](#)
- [28] PUHALSKII, A. (1991). On functional principle of large deviations. In *New Trends in Probability and Statistics* **1** 198–219. VSP, Utrecht. [MR1200917](#)
- [29] SOWERS, R. (1992). Large deviations for a reaction–diffusion equation with non-Gaussian perturbations. *Ann. Probab.* **20** 504–537. [MR1143433](#)
- [30] SPOHN, H. (1991). *Large Scale Dynamics of Interacting Particles*. Springer, Berlin.
- [31] TATARU, D. (1992). Viscosity solutions of Hamilton–Jacobi equations with unbounded nonlinear terms. *J. Math. Anal. Appl.* **163** 345–392. [MR1145836](#)
- [32] TATARU, D. (1994). Viscosity solutions for Hamilton–Jacobi equations with unbounded nonlinear term: A simplified approach. *J. Differential Equations* **111** 123–146. [MR1280618](#)

DEPARTMENT OF MATHEMATICS AND STATISTICS  
UNIVERSITY OF MASSACHUSETTS–AMHERST  
AMHERST, MASSACHUSETTS 01003  
USA  
E-MAIL: [feng@math.umass.edu](mailto:feng@math.umass.edu)